2

Metric Spaces

What is the minimum of structure one needs to have on a set in order to be able to speak of continuity?

If f is a function defined on a subset of $\mathbb{R}$—or, more generally, of Euclidean n-space $\mathbb{R}^n$—we say that f is continuous at x_0 if "$f(x)$ approaches $f(x_0)$ as x approaches x_0." With ϵ and δ, this statement can be made sufficiently precise for mathematical purposes.

> *For each $\epsilon > 0$, there is $\delta > 0$ such that $|f(x) - f(x_0)| < \epsilon$ for all x such that $|x - x_0| < \delta$.*

Crucial for the definition of continuity thus seems to be that we can measure the distance between two real numbers (or, rather, two vectors in Euclidean n-space).

If we want to speak of continuity of functions defined on more general sets, we should thus have a meaningful way to speak of the distance between two points of such a set: this, in a nutshell, is the idea behind a metric space.

2.1 Definitions and Examples

In Euclidean 2-space, the distance between two points (x_1, x_2) and (y_1, y_2) is defined as $\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$. More generally, in Euclidean n-space $\mathbb{R}^n$, one defines, for $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, their distance as

$$d(x, y) := \sqrt{\sum_{j=1}^{n} (x_j - y_j)^2}.$$

The Euclidean distance has the following properties.

1. $d(x, y) \geq 0$ for all $x, y \in \mathbb{R}^n$ with $d(x, y) = 0$ if and only if $x = y$;
2. $d(x, y) = d(y, x)$ for all $x, y \in \mathbb{R}^n$;
3. $d(x, z) \leq d(x, y) + d(y, z)$ for $x, y, z \in \mathbb{R}^n$.

In the definition of a metric space, these three properties of the Euclidean distance are axiomatized.

Definition 2.1.1. *Let X be a set. A* metric *on X is a map $d\colon X \times X \to \mathbb{R}$ with the following properties:*

(a) $d(x, y) \geq 0$ *for all* $x, y \in X$ *with* $d(x, y) = 0$ *if and only if* $x = y$ (positive definiteness)*;*
(b) $d(x, y) = d(y, x)$ *for all* $x, y \in X$ (symmetry)*;*
(c) $d(x, z) \leq d(x, y) + d(y, z)$ *for* $x, y, z \in X$ (triangle inequality).

A set together with a metric is called a metric space.

We often denote a metric space X whose metric is d by (X, d); sometimes, if the metric is obvious or irrelevant, we may also simply write X.

Examples 2.1.2. (a) $\mathbb{R}^n$ with the Euclidean distance is a metric space.
(b) Let (X, d) be a metric space, and let Y be a subset of X. Then the restriction of d to $Y \times Y$ turns Y into a metric space of its own. The metric space $(Y, d|_{Y\times Y})$ is called a *subspace* of X. In particular, any subset of $\mathbb{R}^n$ equipped with the Euclidean distance is a subspace of $\mathbb{R}^n$.
(c) Let E be a linear space (over $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$). A *norm* on E is a map $\|\cdot\|\colon E \to \mathbb{R}$ such that: (i) $\|x\| \geq 0$ for all $x \in E$ with $\|x\| = 0$ if and only if $x = 0$; (ii) $\|\lambda x\| = |\lambda|\|x\|$ for $\lambda \in \mathbb{F}$ and $x \in E$; (iii) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in E$ (a linear space equipped with a norm is called a *normed space*). For $x, y \in E$, define

$$d(x, y) := \|x - y\|.$$

This turns E into a metric space. For example, let E be $C([0, 1], \mathbb{F})$, the space of all continuous $\mathbb{F}$-valued functions on $[0, 1]$. Then there are several norms on E, for example, $\|\cdot\|_1$ defined by

$$\|f\|_1 := \int_0^1 |f(t)|\, dt \qquad (f \in E)$$

or $\|\cdot\|_\infty$ given by

$$\|f\|_\infty := \sup\{|f(t)| : t \in [0, 1]\} \qquad (f \in E).$$

Each of them turns E into a normed space.
(d) Let $S \neq \varnothing$ be a set, and let (Y, d) be a metric space. A function $f\colon S \to Y$ is said to be *bounded* if

$$\sup_{x,y\in S} d(f(x), f(y)) < \infty$$

The set

$$B(S, Y) := \{f\colon S \to Y : f \text{ is bounded}\}$$

becomes a metric space through D defined by

$$D(f, g) := \sup_{x\in S} d(f(x), g(x)) \qquad (f, g \in B(S, Y)).$$

(e) France is a centralized country: every train that goes from one French city to another has to pass through Paris. This is slightly exaggerated, but not too much, as the map shows.

Fig. 2.1: Map of the French railroad network

This motivates the name *French railroad metric* for the following construction. Let (X, d) be a metric space ("France"), and fix $p \in X$ ("Paris"). Define a new metric d_p on X by letting

$$d_p(x, y) := \begin{cases} 0, & x = y, \\ d(x, p) + d(p, y), & \text{otherwise,} \end{cases}$$

for $x, y \in X$. Then (X, d_p) is again a metric space.

(f) Let (X, d) be any metric space, and define $\tilde{d}\colon X \times X \to \mathbb{R}$ via

$$\tilde{d}(x,y) := \frac{d(x,y)}{1+d(x,y)} \qquad (x, y \in X).$$

We claim that $\tilde{d}$ is a metric on X. It is obvious that $\tilde{d}$ is positive definite and symmetric. Hence, all we have to show is that the triangle inequality holds. First note that the function

$$[0,\infty) \to \mathbb{R}, \quad t \mapsto \frac{t}{1+t} \tag{$*$}$$

is increasing (this can be verified, for instance, through differentiation). Let $x, y, z \in X$, and observe that

$$\begin{aligned}
\tilde{d}(x,z) &= \frac{d(x,z)}{1+d(x,z)} \\
&\leq \frac{d(x,y)+d(y,z)}{1+d(x,y)+d(y,z)}, \qquad \text{because } (*) \text{ is increasing}, \\
&= \frac{d(x,y)}{1+d(x,y)+d(y,z)} + \frac{d(y,z)}{1+d(x,y)+d(y,z)} \\
&\leq \frac{d(x,y)}{1+d(x,y)} + \frac{d(y,z)}{1+d(y,z)} \\
&= \tilde{d}(x,y) + \tilde{d}(y,z).
\end{aligned}$$

Consequently, $\tilde{d}$ is indeed a metric on X.

(g) A *semimetric* d on a set X satisfies the same axioms as a metric with one exception: it is possible for $x, y \in X$ with $x \neq y$ that $d(x,y) = 0$. If d is a semimetric, then $\tilde{d}$ as constructed in the previous example is also a semimetric. Let X be equipped with a sequence $(d_n)_{n=1}^\infty$ of semimetrics such that, for any $x, y \in X$ with $x \neq y$, there is $n \in \mathbb{N}$ with $d_n(x,y) > 0$. Then $d\colon X \times X \to \mathbb{R}$ defined by

$$d(x,y) := \sum_{n=1}^\infty \frac{1}{2^n} \frac{d_n(x,y)}{1+d_n(x,y)} \qquad (x, y \in X)$$

is a metric. Clearly, d is symmetric and satisfies the triangle inequality, and if $x, y \in X$ are such that $x \neq y$, there is $n \in \mathbb{N}$ with $d_n(x,y) > 0$, so that $d(x,y) \geq \frac{1}{2^n} \frac{d_n(x,y)}{1+d_n(x,y)} > 0$.

(h) The previous example can be used, for instance, to turn a Cartesian product X of countably many metric spaces $((X_n, d_n))_{n=1}^\infty$ into a metric space again. For each $n \in \mathbb{N}$, the map

$$\delta_n\colon X \times X \to [0,\infty), \quad ((x_1, x_2, x_3, \ldots), (y_1, y_2, y_3, \ldots)) \mapsto d_n(x_n, y_n)$$

is a semimetric. Moreover, if $x = (x_1, x_2, x_3, \ldots)$ and $y = (y_1, y_2, y_3, \ldots)$ are different points of X, there is at least one coordinate $n \in \mathbb{N}$ such that $x_n \neq y_n$, so that $\delta_n(x, y) = d_n(x_n, y_n) > 0$. For $x = (x_1, x_2, x_3, \ldots)$ and $y = (y_1, y_2, y_3, \ldots)$ in X, let

$$d(x,y) := \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\delta_n(x,y)}{1 + \delta_n(x,y)} = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{d_n(x_n,y_n)}{1 + d_n(x_n,y_n)}.$$

Then d is a metric on X.

(i) Let X be any set. For $x, y \in X$ define

$$d(x,y) := \begin{cases} 0, & x = y, \\ 1, & \text{otherwise.} \end{cases}$$

Then (X, d) is easily seen to be a metric space. (Metric spaces of this form are called *discrete*.)

Exercises

1. Let S be any set, and let X consist of the finite subsets of S. Show that

$$d \colon X \times X \to [0, \infty), \quad (A, B) \mapsto |(A \setminus B) \cup (B \setminus A)|$$

is a metric on X.

2. Verify Example 2.1.2(d) in detail.
3. Let $S \neq \varnothing$ be a set, and let E be a normed space. Show that

$$\|f\|_\infty := \sup\{\|f(x)\| : x \in S\} \qquad (f \in B(S,E))$$

defines a norm on $B(S, E)$. How does $\|\cdot\|_\infty$ relate to the metric D from the previous exercise?

4. Let $(E, \|\cdot\|)$ be a normed space, and define $|||\cdot||| \colon E \to [0, \infty)$ by letting

$$|||x||| := \frac{\|x\|}{1 + \|x\|} \qquad (x \in E).$$

Is $|||\cdot|||$ a norm on E?

5. Let X be any set, and let $d \colon X \times X \to [0, \infty)$ be a semimetric. For $x, y \in X$, define $x \approx y$ if and only if $d(x, y) = 0$.
   (a) Show that $\approx$ is an equivalence relation on X.
   (b) For $x \in X$, let $[x]$ denote its equivalence class with respect to $\approx$, and let $X/\!\approx$ denote the collection of all $[x]$ with $x \in X$. Show that

$$(X/\!\approx) \times (X/\!\approx) \to [0, \infty), \quad ([x],[y]) \mapsto d(x,y)$$

defines a metric on $X/\!\approx$.

2.2 Open and Closed Sets

We start with the definition of an open ball in a metric space:

Definition 2.2.1. *Let (X, d) be a metric space, let $x_0 \in X$, and let $r > 0$. The* open ball *centered at x_0 with radius r is defined as*

$$B_r(x_0) := \{x \in X : d(x, x_0) < r\}.$$

Of course, in Euclidean 2- or 3-space, this definition coincides with the usual intuitive one. Nevertheless, even though open balls are defined with the intuitive notions of Euclidean space in mind, matters can turn out to be surprisingly counterintuitive:

Examples 2.2.2. (a) Let (X, d) be a discrete metric space, let $x_0 \in X$, and let $r > 0$. Then

$$B_r(x_0) = \begin{cases} \{x_0\}, & r < 1, \\ X, & r \geq 1, \end{cases}$$

holds; that is, each open ball is a singleton subset or the whole space.

(b) Let (X, d) be any metric space, let $p \in X$, and let d_p be the corresponding French railroad metric. To tell open balls in (X, d) and (X, d_p) apart, we write $B_r(x_0; d)$ and $B_r(x_0; d_p)$, respectively, for $x_0 \in X$ and $r > 0$. Let $x_0 \in X$, and let $r > 0$. Since, for $x \in X$ with $x \neq x_0$, we have

$$d_p(x, x_0) = d(x, p) + d(p, x_0) < r \quad \iff \quad d(x, p) < r - d(p, x_0),$$

the following dichotomy holds.

$$B_r(x_0; d_p) = \begin{cases} \{x_0\}, & \text{if } r \leq d(p, x_0), \\ B_{r - d(p, x_0)}(p; d) \cup \{x_0\}, & \text{otherwise.} \end{cases}$$

Like the notion of an open ball, the notion of an open set extends from Euclidean space to arbitrary metric spaces.

Definition 2.2.3. *Let (X, d) be a metric space. A set $U \subset X$ is called* open *if, for each $x \in U$, there is $\epsilon > 0$ such that $B_\epsilon(x) \subset U$.*

If our choice of terminology is to make any sense, an open ball in a metric space better be an open set. Indeed, this is true.

Example 2.2.4. Let (X, d) be a metric space, let $x_0 \in X$, and let $r > 0$. For $x \in B_r(x_0)$, choose $\epsilon := r - d(x, x_0) > 0$. Hence, we have for $y \in B_\epsilon(x)$:

$$d(y, x_0) \leq d(y, x) + d(x, x_0) < \epsilon + d(x, x_0) = r - d(x, x_0) + d(x, x_0) = r.$$

It follows that $B_\epsilon(x) \subset B_r(x_0)$.

The following proposition lists the fundamental properties of open sets.

Proposition 2.2.5. *Let (X, d) be a metric space. Then:*

(i) *$\varnothing$ and X are open;*
(ii) *If $\mathcal{U}$ is a family of open subsets of X, then $\bigcup\{U : U \in \mathcal{U}\}$ is open;*
(iii) *If U_1 and U_2 are open subsets of X, then $U_1 \cap U_2$ is open.*

Proof. (i) is clear.

For (ii), let $\mathcal{U}$ be a family of open sets in X, and let $x \in \bigcup\{U : U \in \mathcal{U}\}$. Then there is $U_0 \in \mathcal{U}$ with $x \in U_0$, and since U_0 is open there is $\epsilon > 0$ such that

$$B_\epsilon(x) \subset U_0 \subset \bigcup\{U : U \in \mathcal{U}\}.$$

Hence, $\bigcup\{U : U \in \mathcal{U}\}$ is open.

Let $U_1, U_2 \subset X$ be open, and let $x \in U_1 \cap U_2$. Since U_1 and U_2 are open, there are $\epsilon_1, \epsilon_2 > 0$ such that $B_{\epsilon_j}(x) \subset U_j$ for $j = 1, 2$. Let $\epsilon := \min\{\epsilon_1, \epsilon_2\}$. Then it is immediate that $B_\epsilon(x) \subset U_1 \cap U_2$. This proves (iii). □

Proposition 2.2.5(i) may seem odd at the first glance. The closed unit interval in $\mathbb{R}$ is a subspace of $\mathbb{R}$, thus a metric space in its own right, and thus open by Proposition 2.2.5(i). But, of course, we know that $[0, 1]$ is *not* open. How is this possible? The answer is that openness (as well as all the notions that are derived from it) depends on the context of a given metric space. Thus, $[0, 1]$ is open in $[0, 1]$, but not open in $\mathbb{R}$.

Example 2.2.6. Let (X, d) be a discrete metric space, and let $S \subset X$. Then

$$S = \bigcup_{x \in S} \{x\} = \bigcup_{x \in S} B_1(x)$$

is open; that is, all subsets of X are open.

A notion closely related to open sets is that of a neighborhood of a point.

Definition 2.2.7. *Let (X, d) be a metric space, and let $x \in X$. A subset N of X is called a* neighborhood *of x if there is an open subset U of X with $x \in U \subset N$. The collection of all neighborhoods of x is denoted by $\mathcal{N}_x$.*

Proposition 2.2.8. *Let (X, d) be a metric space, and let $x \in X$. Then:*

(i) *A subset N of X belongs to $\mathcal{N}_x$ if and only if there is $\epsilon > 0$ such that $B_\epsilon(x) \subset N$;*
(ii) *If $N \in \mathcal{N}_x$ and $M \supset N$, then $M \in \mathcal{N}_x$;*
(iii) *If $N_1, N_2 \in \mathcal{N}_x$, then $N_1 \cap N_2 \in \mathcal{N}_x$.*

Moreover, a subset U of X is open if and only if $U \in \mathcal{N}_y$ for each $y \in U$.

Proof. Suppose that $N \subset X$ is such that there is $\epsilon > 0$ such that $B_\epsilon(x) \subset N$. Since $B_\epsilon(x)$ is open, it follows that $N \in \mathcal{N}_x$. Conversely, suppose that $N \in \mathcal{N}_x$. Then there is an open subset U of N with $x \in U$. By the definition of openness, there is $\epsilon > 0$ such that $B_\epsilon(x) \subset U \subset N$. This proves (i).

(ii) is obvious, and (iii) follows immediately from Proposition 2.2.5(iii).

Let $U \subset X$ be open. Then, clearly, U is a neighborhood of each of its points. Conversely, let $U \subset X$ be any set with that property. By the definition of a neighborhood, there is, for each $y \in U$, an open subset U_y of U with $y \in U_y$. Since $U = \bigcup_{y \in U} U_y$, Proposition 2.2.5(ii) yields that U is open. $\square$

As in Euclidean space, we define a set to be closed if its complement is open.

Definition 2.2.9. *Let (X, d) be a metric space. A subset F of X is called* closed *if $X \setminus F$ is open.*

Examples 2.2.10. (a) Let (X, d) be any metric space, let $x_0 \in X$, and let $r > 0$. The *closed ball* centered at x_0 with radius r is defined as

$$B_r[x_0] := \{x \in X : d(x, x_0) \leq r\}.$$

We claim that $B_r[x_0]$ is indeed closed. To show this, let $x \in X \setminus B_r[x_0]$, that is, such that $d(x, x_0) > r$. Let $\epsilon := d(x, x_0) - r > 0$, and let $y \in B_\epsilon(x)$. Since $d(x, x_0) \leq d(x, y) + d(y, x_0)$, we obtain that

$$d(y, x_0) \geq d(x, x_0) - d(x, y) > d(x, x_0) - \epsilon = d(x, x_0) - (d(x, x_0) - r) = r.$$

It follows that $B_\epsilon(x) \subset X \setminus B_r[x_0]$. Consequently, $X \setminus B_r[x_0]$ is open and $B_r[x_0]$ is closed.

(b) In a discrete metric space, every subset is both open and closed.

The following is a straightforward consequence of Proposition 2.2.5.

Proposition 2.2.11. *Let (X, d) be a metric space. Then:*

(i) *$\varnothing$ and X are closed;*
(ii) *If $\mathcal{F}$ is a family of closed subsets of X, then $\bigcap\{F : F \in \mathcal{F}\}$ is closed;*
(iii) *If F_1 and F_2 are closed subsets of X, then $F_1 \cup F_2$ is closed.*

Of course, in most metric spaces there are many sets that are neither open nor closed. Nevertheless, we can make the following definition.

Definition 2.2.12. *Let (X, d) be a metric space. For each $S \subset X$, the* closure *of S is defined as*

$$\overline{S} := \bigcap\{F : F \subset X \text{ is closed and contains } S\}.$$

From Proposition 2.2.11(ii) it is immediate that the closure of a set is a closed set. The following is an alternative description of the closure.

Proposition 2.2.13. *Let (X, d) be a metric space, and let $S \subset X$. Then we have:*

$$\begin{aligned}\overline{S} &= \{x \in X : N \cap S \neq \varnothing \text{ for all } N \in \mathcal{N}_x\} \\ &= \{x \in X : B_\epsilon(x) \cap S \neq \varnothing \text{ for all } \epsilon > 0\}.\end{aligned}$$

Proof. Each open ball is a neighborhood of its center, and any neighborhood of a point contains an open ball centered at that point; therefore

$$\{x \in X : N \cap S \neq \varnothing \text{ for all } N \in \mathcal{N}_x\} = \{x \in X : B_\epsilon(x) \cap S \neq \varnothing \text{ for all } \epsilon > 0\}$$

holds. We denote this set by $\mathrm{cl}(S)$.

Let $x \in \overline{S}$, and let $N \in \mathcal{N}_x$. Then there is an open subset U of X contained in N with $x \in U$. Assume that $N \cap S = \varnothing$, so that $U \cap S = \varnothing$ (i.e., $S \subset X \setminus U$). Since $X \setminus U$ is closed, it follows that $\overline{S} \subset X \setminus U$ and thus $x \in X \setminus U$, which is a contradiction. Consequently, $x \in \mathrm{cl}(S)$ holds.

Conversely, let $x \in \mathrm{cl}(S)$, and assume that $x \notin \overline{S}$. Then $U := X \setminus \overline{S}$ is an open set containing x (thus belonging to $\mathcal{N}_x$) having empty intersection with S. This contradicts $x \in \mathrm{cl}(S)$. □

Examples 2.2.14. (a) Any open interval in $\mathbb{R}$ contains a rational number. Hence, we have $\overline{\mathbb{Q}} = \mathbb{R}$.

(b) Let (X, d) be any metric space. It is obvious that $\overline{B_r(x_0)} \subset B_r[x_0]$ for all $x_0 \in X$ and $r > 0$. In general, equality need not hold. If (X, d) is discrete and has more than one element, we have for any $x_0 \in X$ that

$$\overline{B_1(x_0)} = \overline{\{x_0\}} = \{x_0\} \subsetneq X = B_1[x_0].$$

(c) Let E be a normed space, let $x_0 \in E$, and let $r > 0$. We claim that (in this particular situation) $\overline{B_r(x_0)} = B_r[x_0]$ holds. In view of the previous example, only $B_r[x_0] \subset \overline{B_r(x_0)}$ needs proof. Let $x \in B_r[x_0]$, and let $\epsilon > 0$. Choose $\delta \in (0, 1)$ such that $\delta \|x - x_0\| < \epsilon$, and let

$$y := x_0 + (1 - \delta)(x - x_0) = (1 - \delta)x + \delta x_0,$$

so that

$$\|y - x_0\| = (1 - \delta)\|x - x_0\| \leq (1 - \delta)r < r;$$

that is, $y \in B_r(x_0)$. Furthermore, we have

$$\|y - x\| = \|(1 - \delta)x + \delta x_0 - x\| = \delta\|x - x_0\| < \epsilon,$$

and thus $y \in B_\epsilon(x)$. From Proposition 2.2.13, we conclude that $x \in \overline{B_r(x_0)}$.

The closure of a set is important in connection with two further topological concepts: density and the boundary.

Definition 2.2.15. *Let (X, d) be a metric space.*

(a) *A subset D of X is said to be* dense *in X if $\overline{D} = X$.*
(b) *If X has a dense countable subset, then X is called* separable.

Examples 2.2.16. (a) $\mathbb{Q}$ is dense in $\mathbb{R}$. In particular, $\mathbb{R}$ is separable.

(b) A subset S of a discrete metric space (X, d) is dense if and only if $S = X$. In particular, X is separable if and only if it is countable.

The following hereditary property of separability is somewhat surprising, but very useful.

Theorem 2.2.17. *Let (X, d) be a separable metric space, and let Y be a subspace of X. Then Y is also separable.*

Proof. Let $C = \{x_1, x_2, x_3, \ldots\}$ be a dense countable subset of X. One might be tempted to use $Y \cap C$ as a dense (and certainly countable) subset of Y, but this may not work: if $X \neq C$, take $Y = X \setminus C$, for example.

Let

$$\mathbb{A} := \left\{ (n, m) \in \mathbb{N} \times \mathbb{N} : \text{there is } y \in Y \text{ such that } d(y, x_n) < \frac{1}{m} \right\}.$$

For each $(n, m) \in \mathbb{A}$, choose $y_{n,m} \in Y$ with $d(y_{n,m}, x_n) < \frac{1}{m}$. Then $C_Y := \{y_{n,m} : (n, m) \in \mathbb{A}\}$ is a countable subset of Y. We claim that C_Y is also dense in Y. Let $y \in Y$, and let $\epsilon > 0$. Choose $m \in \mathbb{N}$ such that $\frac{1}{m} \leq \frac{\epsilon}{2}$. Since C is dense in X, there is $n \in \mathbb{N}$ such that $d(y, x_n) < \frac{1}{m}$. By the definition of $\mathbb{A}$, this means that $(n, m) \in \mathbb{A}$. It follows that

$$d(y, y_{n,m}) \leq d(y, x_n) + d(x_n, y_{n,m}) < \frac{2}{m} \leq \epsilon.$$

By Proposition 2.2.13, this means that y lies in the closure of C_Y in Y. $\square$

Examples 2.2.18. (a) The irrational numbers are a separable subspace of $\mathbb{R}$.

(b) Let $X = B(\mathbb{N}, \mathbb{R})$ be equipped with the metric introduced in Example 2.1.2(d); that is,

$$d(f, g) = \sup_{n \in \mathbb{N}} |f(n) - g(n)| \qquad (f, g \in X).$$

We claim that X is not separable. We assume towards a contradiction that X *is* separable. Let Y denote the subspace of X consisting of all $\{0, 1\}$ valued functions. From Theorem 2.2.17, it follows that Y is separable, too. Since, for $f, g \in Y$, we have

$$d(f, g) = \sup_{n \in \mathbb{N}} |f(n) - g(n)| = \begin{cases} 0, & f = g, \\ 1, & f \neq g, \end{cases}$$

it follows that Y is a discrete metric space and therefore must be countable. However, the map

$$Y \to [0, 1], \quad f \mapsto \sum_{n=1}^{\infty} \frac{f(n)}{2^n}$$

is surjective, and $[0, 1]$ is not countable. This is a contradiction.

To motivate the notion of boundary, we first consider an example.

Example 2.2.19. Let $(E, \|\cdot\|)$ be a normed space, let $x_0 \in E$, and let $r > 0$. Then, intuitively, one might view the boundary of the open ball $B_r(x_0)$ as the *sphere*

$$S_r[x_0] := \{x \in E : \|x - x_0\| = r\}.$$

Let $x \in S_r[x_0]$, and let $\epsilon > 0$. Let $\delta \in (0,1)$ be such that $\delta\|x - x_0\| < \epsilon$, and let $y := x_0 + (1-\delta)(x - x_0)$. As in Example 2.2.14(c), it follows that $y \in B_\epsilon(x) \cap B_r(x_0)$, so that

$$B_\epsilon(x) \cap B_r(x_0) \neq \varnothing \qquad \text{and} \qquad B_\epsilon(x) \cap (E \setminus B_r(x_0)) \neq \varnothing. \qquad (**)$$

On the other hand, since $B_r(x_0)$ and $E \setminus B_r[x_0]$ are open, it follows that any element x of E satisfying $(**)$ for each $\epsilon > 0$ must lie in $S_r[x_0]$.

In view of this example, we define the following.

Definition 2.2.20. *Let (X, d) be a metric space, and let $S \subset X$. Then the* boundary *of S is defined as*

$$\partial S := \{x \in X : B_\epsilon(x) \cap S \neq \varnothing \text{ and } B_\epsilon(x) \cap (X \setminus S) \neq \varnothing \text{ for all } \epsilon > 0\}.$$

An argument similar to that at the beginning of the proof of Proposition 2.2.13 yields immediately that

$$\partial S = \{x \in X : N \cap S \neq \varnothing \text{ and } N \cap (X \setminus S) \neq \varnothing \text{ for all } N \in \mathcal{N}_x\}$$

for each subset S of a metric space X.

Proposition 2.2.21. *Let (X, d) be a metric space, and let $S \subset X$. Then:*

(i) $\partial S = \partial(X \setminus S)$;
(ii) ∂S *is closed;*
(iii) $\overline{S} = S \cup \partial S$.

Proof. (i) is a triviality.

For (ii), let $x \in X \setminus \partial S$; that is, there is $N \in \mathcal{N}_x$ such that $N \cap S = \varnothing$ or $N \cap (X \setminus S) = \varnothing$. Let $U \subset N$ be open such that $x \in U$. It follows that $U \cap S = \varnothing$ or $U \cap (X \setminus S) = \varnothing$. Since U is a neighborhood of each of its points, it follows that $U \subset X \setminus \partial S$. Hence, $X \setminus \partial S$ is a neighborhood of x. Since x was arbitrary, it follows that $X \setminus \partial S$ is open.

For (iii), note that, by Proposition 2.2.13, $\partial S \subset \overline{S}$ holds, so that $S \cup \partial S \subset \overline{S}$. Conversely, let $x \in \overline{S}$, and suppose that $x \notin S$. For each $N \in \mathcal{N}_x$, it is clear that $N \cap (X \setminus S) \neq \varnothing$, and Proposition 2.2.13 yields that $N \cap S \neq \varnothing$ as well. $\square$

The closure of a given set is, by definition, the smallest closed set containing it. Analogously, one defines the largest open set contained in a given set.

Definition 2.2.22. *Let (X,d) be a metric space. For each $S \subset X$, the* interior *of S is defined as*

$$\overset{\circ}{S} := \bigcup \{U : U \subset X \text{ is open and contained in } S\}.$$

The following proposition characterizes the interior of a set:

Proposition 2.2.23. *Let (X,d) be a metric space, and let $S \subset X$. Then we have:*

$$\overset{\circ}{S} = \{x \in X : S \in \mathcal{N}_x\} = S \setminus \partial S.$$

Proof. Let $x \in \overset{\circ}{S}$. Then there is an open subset U of S with $x \in U$, so that $S \in \mathcal{N}_x$. Conversely, if $S \in \mathcal{N}_x$, then there is an open set U of X with $x \in U \subset S$, so that $x \in \overset{\circ}{S}$.

Let $x \in \overset{\circ}{S}$, so that $S \in \mathcal{N}_x$ by the foregoing. Since, trivially, $S \cap (X \setminus S) = \varnothing$, we see that $x \notin \partial S$. Conversely, let $x \in S \setminus \partial S$. Then there is $N \in \mathcal{N}_x$ such that $N \cap (X \setminus S) = \varnothing$. Let $U \subset N$ be open in X such that $x \in U$. It follows that $U \cap (X \setminus S) = \varnothing$ and therefore $U \subset S$. Consequently, $x \in U \subset \overset{\circ}{S}$ holds. □

Exercises

1. Show that a finite subset of a metric space is closed.
2. Let $(E, \|\cdot\|)$ be a normed space, let $U \subset E$ be open, and let $S \subset E$ be any set. Show that $S + U := \{x + y : x \in S,\ y \in U\}$ is open in E.
3. Let $U \subset \mathbb{R}$ be open.
   (a) For each $x \in U$, let I_x be the union of all open intervals contained in U and containing x. Show that I_x is an open (possibly unbounded) interval.
   (b) For $x, y \in U$, show that $I_x = I_y$ or $I_x \cap I_y = \varnothing$.
   (c) Conclude that U is a union of countably many, pairwise disjoint open intervals.
4. Let (X,d) be a metric space, and let $S \subset X$. The *distance* of $x \in X$ to S is defined as
$$\operatorname{dist}(x,S) := \inf\{d(x,y) : y \in S\}$$
(where $\operatorname{dist}(x,S) = \infty$ if $S = \varnothing$). Show that $\overline{S} = \{x \in X : \operatorname{dist}(x,S) = 0\}$.
5. Let Y be the subspace of $B(\mathbb{N}, \mathbb{F})$ consisting of those sequences tending to zero. Show that Y is separable.
6. Let (X,d) be a metric space, and let Y be a subspace of X. Show that $U \subset Y$ is open in Y if and only if there is $V \subset X$ that is open in X such that $U = Y \cap V$.

2.3 Convergence and Continuity

The notion of convergence in $\mathbb{R}^n$ carries over to metric spaces almost verbatim.

Definition 2.3.1. *Let (X, d) be a metric space. A sequence $(x_n)_{n=1}^{\infty}$ in X is said to* converge *to $x \in X$ if, for each $\epsilon > 0$, there is $n_\epsilon \in \mathbb{N}$ such that $d(x_n, x) < \epsilon$ for all $n \geq n_\epsilon$. We then say that x is the* limit *of $(x_n)_{n=1}^{\infty}$ and write $x = \lim_{n\to\infty} x_n$ or $x_n \to x$.*

It is straightforward to verify that a sequence $(x_n)_{n=1}^{\infty}$ in a metric space converges to x if and only if, for each $N \in \mathcal{N}_x$, there is $n_N \in \mathbb{N}$ such that $x_n \in N$ for all $n \geq n_N$.

Examples 2.3.2. (a) Let (X, d) be a discrete metric space, and let $(x_n)_{n=1}^{\infty}$ be a sequence in X that converges to $x \in X$. Then there is $n_1 \in \mathbb{N}$ such that $d(x_n, x) < 1$ for $n \geq n_1$; that is, $x_n = x$ for $n \geq n_1$. Hence, every convergent sequence in a discrete metric space is eventually constant.

(b) Let $C([0, 1], \mathbb{F})$ be equipped with the metric induced by $\|\cdot\|_\infty$ (Example 2.1.2(c)). We claim that a sequence $(f_n)_{n=1}^{\infty}$ in $C([0, 1], \mathbb{F})$ converges to $f \in C([0, 1], \mathbb{F})$ with respect to that metric if and only if it converges (to f) uniformly on $[0, 1]$. Suppose first that $\|f_n - f\|_\infty \to 0$, and let $\epsilon > 0$. Then there is $n_\epsilon \in \mathbb{N}$ such that

$$|f_n(t) - f(t)| \leq \|f_n - f\|_\infty < \epsilon \qquad (n \geq n_\epsilon,\ t \in [0, 1]),$$

so that $f_n \to f$ uniformly on $[0, 1]$. Conversely, let $(f_n)_{n=1}^{\infty}$ converge to f uniformly on $[0, 1]$, and let $\epsilon > 0$. By the definition of uniform convergence, there is $n_\epsilon \in \mathbb{N}$ such that

$$|f_n(t) - f(t)| < \frac{\epsilon}{2} \qquad (n \geq n_\epsilon,\ t \in [0, 1])$$

and consequently,

$$\|f_n - f\|_\infty = \sup\{|f_n(t) - f(t)| : t \in [0, 1]\} \leq \frac{\epsilon}{2} < \epsilon \qquad (n \geq n_\epsilon).$$

Hence, we have convergence with respect to $\|\cdot\|_\infty$.

As in $\mathbb{R}^n$, the limit of a sequence in a metric space is unique.

Proposition 2.3.3. *Let (X, d) be a metric space, let $(x_n)_{m=1}^{\infty}$ be a sequence in X, and let $x, x' \in X$ be such that $(x_n)_{n=1}^{\infty}$ converges to both x and x'. Then x and x' are equal.*

Proof. Assume that $x \neq x'$, so that $\epsilon := \frac{1}{2} d(x, x') > 0$. Since $x_n \to x$, there is $n_1 \in \mathbb{N}$ such that $d(x_n, x) < \epsilon$ for $n \geq n_1$, and since $x_n \to x'$, too, there is $n_2 \in \mathbb{N}$ such that $d(x_n, x') < \epsilon$ for $n \geq n_2$. Let $n := \max\{n_1, n_2\}$, so that

$$d(x, x') \leq d(x, x_n) + d(x_n, x') < \epsilon + \epsilon = d(x, x'),$$

which is nonsense. $\square$

Here is the idea of the proof of Proposition 2.3.3 in a sketch.

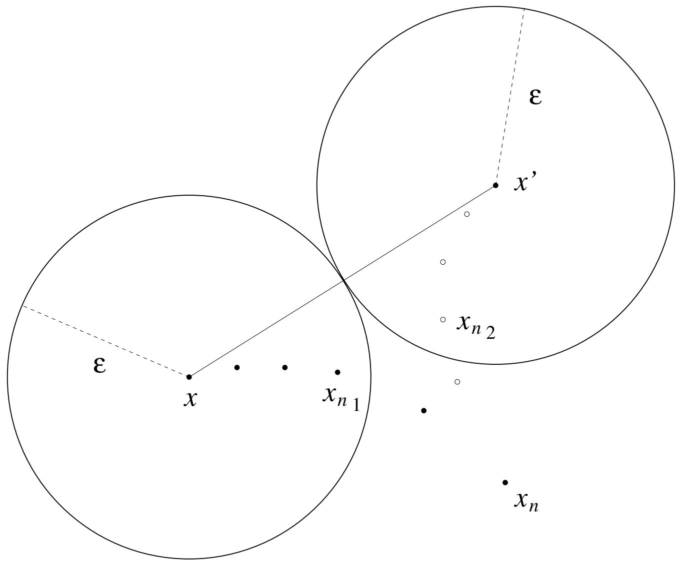

Fig. 2.2: Uniqueness of the limit

Also, as in $\mathbb{R}^n$, convergence in metric spaces can be used to characterize the closed subsets.

Proposition 2.3.4. *Let (X, d) be a metric space, and let $S \subset X$. Then $\overline{S}$ consists of those points in X that are the limit of a sequence in S.*

Proof. Let $x \in X$ be the limit of a sequence $(x_n)_{n=1}^\infty$ in S, and let $\epsilon > 0$. By the definition of convergence, there is $n_\epsilon \in \mathbb{N}$ such that $d(x_n, x) < \epsilon$ for $n \geq n_\epsilon$; that is, $x_n \in B_\epsilon(x)$ for $n \geq n_\epsilon$. In particular, $B_\epsilon(x) \cap S$ is nonempty. Since $\epsilon > 0$ is arbitrary, it follows that $x \in \overline{S}$ by Proposition 2.2.13.

Conversely, let $x \in \overline{S}$. By Proposition 2.2.13, we have $B_{\frac{1}{n}}(x) \cap S \neq \varnothing$ for each $n \in \mathbb{N}$; there is thus, for each $n \in \mathbb{N}$, some $x_n \in S$ with $d(x_n, x) < \frac{1}{n}$. It is clear that the sequence $(x_n)_{n=1}^\infty$ converges to x. $\square$

Corollary 2.3.5. *Let (X, d) be a metric space. Then $F \subset X$ is closed if and only if every sequence in F that converges in X has its limit in F.*

Of course, with a notion of convergence at hand, continuity of functions can be defined.

Definition 2.3.6. *Let (X, d_X) and (Y, d_Y) be metric spaces, and let $x_0 \in X$. Then $f : X \to Y$ is said to be* continuous at x_0 *if, for each sequence $(x_n)_{n=1}^\infty$ in X that converges to x_0, we have $\lim_{n\to\infty} f(x_n) = f(x_0)$.*

The following characterization holds.

Theorem 2.3.7. *Let (X, d_X) and (Y, d_Y) be metric spaces, and let $x_0 \in X$. Then the following are equivalent for $f: X \to Y$.*

(i) *f is continuous at x_0.*
(ii) *For each $\epsilon > 0$, there is $\delta > 0$ such that $d_Y(f(x), f(x_0)) < \epsilon$ for all $x \in X$ with $d_X(x, x_0) < \delta$.*
(iii) *For each $\epsilon > 0$, there is $\delta > 0$ such that $B_\delta(x_0) \subset f^{-1}(B_\epsilon(f(x_0)))$.*
(iv) *For each $N \in \mathcal{N}_{f(x_0)}$, we have $f^{-1}(N) \in \mathcal{N}_{x_0}$.*

Proof. (i) $\Longrightarrow$ (ii): Assume otherwise; that is, there is $\epsilon_0 > 0$ such that, for each $\delta > 0$, there is $x_\delta \in X$ with $d_X(x_\delta, x_0) < \delta$, but $d_Y(f(x_\delta), f(x_0)) \geq \epsilon_0$. For $n \in \mathbb{N}$, let $x'_n := x_{\frac{1}{n}}$, so that $d(x'_n, x_0) < \frac{1}{n}$ and thus $x'_n \to x_0$. Since, however, $d_Y(f(x'_n), f(x_0)) \geq \epsilon_0$ holds for all $n \in \mathbb{N}$, it is impossible that $f(x'_n) \to f(x_0)$ as required for f to be continuous at x_0.

(iii) is only a rewording of (ii).

(iii) $\Longrightarrow$ (iv): Let $N \in \mathcal{N}_{f(x_0)}$. Hence, there is $\epsilon > 0$ such that $B_\epsilon(x_0) \subset N$. By (iii), there is $\delta > 0$ such that

$$B_\delta(x_0) \subset f^{-1}(B_\epsilon(f(x_0))) \subset f^{-1}(N).$$

This implies that $f^{-1}(N) \in \mathcal{N}_{x_0}$.

(iv) $\Longrightarrow$ (i): Let $(x_n)_{n=1}^\infty$ be a sequence in X with $x_n \to x_0$. Let $N \in \mathcal{N}_{f(x_0)}$, so that $f^{-1}(N) \in \mathcal{N}_{x_0}$. Since $x_n \to x_0$, there is $n_N \in \mathbb{N}$ such that $x_n \in f^{-1}(N)$ for $n \geq n_N$; that is, $f(x_n) \in N$ for $n \geq n_N$. Since $N \in \mathcal{N}_{f(x_0)}$ was arbitrary, this yields $f(x_n) \to f(x_0)$. □

The following definition should also look familiar.

Definition 2.3.8. *Let (X, d_X) and (Y, d_Y) be metric spaces. Then a function $f: X \to Y$ is said to be* continuous *if it is continuous at each point of X.*

Example 2.3.9. Let (X, d) be a metric space. We first claim that

$$|d(x, y) - d(x_0, y_0)| \leq d(x, x_0) + d(y, y_0) \qquad (x, x_0, y, y_0 \in X). \qquad (***)$$

Fix $x, x_0, y, y_0 \in X$, and note that

$$d(x, y) \leq d(x, x_0) + d(x_0, y_0) + d(y_0, y)$$

and therefore

$$d(x, y) - d(x_0, y_0) \leq d(x, x_0) + d(y_0, y).$$

Interchanging the roles of x and x_0 and, respectively, y and y_0, yields $d(x_0, y_0) - d(x, y) \leq d(x, x_0) + d(y_0, y)$. Altogether, we obtain $(***)$. The Cartesian square X^2 becomes a metric space in its own right through

$$\tilde{d}((x, y), (x', y')) := d(x, x') + d(y, y') \qquad ((x, x'), (y, y') \in X^2).$$

The inequality $(***)$ immediately yields that $d: X^2 \to \mathbb{R}$ is continuous if X^2 is equipped with $\tilde{d}$.

Corollary 2.3.10. *Let (X, d_X) and (Y, d_Y) be metric spaces. Then the following are equivalent for $f : X \to Y$.*

(i) *f is continuous.*
(ii) *$f^{-1}(U)$ is open in X for each open subset U of Y.*
(iii) *$f^{-1}(F)$ is closed in X for each closed subset F of Y.*

Proof. (i) $\Longrightarrow$ (ii): Let $U \subset Y$ be open, so that $U \in \mathcal{N}_y$ for each $y \in U$ and thus $U \in \mathcal{N}_{f(x)}$ for each $x \in f^{-1}(U)$. For Theorem 2.3.7(iv), we conclude that $f^{-1}(U) \in \mathcal{N}_x$ for each $x \in f^{-1}(U)$; that is, $f^{-1}(U)$ is a neighborhood of each of its points and thus open.

(ii) $\Longrightarrow$ (iii): Let $F \subset Y$ be closed, so that $Y \setminus F$ is open. Since $X \setminus f^{-1}(F) = f^{-1}(Y \setminus F)$ then must be open by (ii), it follows that $f^{-1}(F)$ is closed. Analogously, (iii) $\Longrightarrow$ (ii) is proved.

(ii) $\Longrightarrow$ (i): If f satisfies (ii), it trivially also satisfies Theorem 2.3.7(iii) for each $x \in X$. $\square$

We now give an example which shows that continuous maps between general metric spaces can be quite different from what we may intuitively expect.

Example 2.3.11. Let (X, d_X) and (Y, d_Y) be metric spaces such that (X, d_X) is discrete, and let $f : X \to Y$ be arbitrary. Let $U \subset Y$ be open. Since in a discrete space every set is open, it follows that $f^{-1}(U)$ is open. Consequently, f must be continuous.

As we have seen, there can be different metrics on one set. For many purposes, it is convenient to view certain metrics as identical.

Definition 2.3.12. *Let X be a set. Two metrics d_1 and d_2 on X are said to be* equivalent *if the identity map on X is continuous both from (X, d_1) to (X, d_2) and from (X, d_2) to (X, d_1).*

In view of Corollary 2.3.10, two metrics d_1 and d_2 on a set X are equivalent if and only if they yield the same open sets (or, equivalently, the same closed sets).

Examples 2.3.13. (a) The Euclidean metric on $\mathbb{R}^n$ and the discrete metric are not equivalent.

(b) For $j = 1, \ldots, n$, let (X_j, d_j) be a metric space. Let $X := X_1 \times \cdots \times X_n$, and for $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in X$, define

$$D_1(x, y) := \sum_{j=1}^{n} d_j(x_j, y_j) \qquad \text{and} \qquad D_\infty(x, y) := \max_{j=1,\ldots,n} d_j(x_j, y_j).$$

Then D_1 and D_∞ are metrics on X satisfying

$$D_\infty(x, y) \leq D_1(x, y) \leq n\, D_\infty(x, y) \qquad (x, y \in X).$$

Consequently, D_1 and D_∞ are equivalent.

(c) Let (X,d) be any metric space, let $p \in X$, and let d_p be the corresponding French railroad metric. Since

$$d(x,y) \leq d_p(x,y) \qquad (x,y \in X),$$

it is easily seen that the identity is continuous from (X,d_p) to (X,d). On the other hand, let $(x_n)_{n=1}^\infty$ be a sequence in X that converges to $x \neq p$ with respect to d. If $x_n \neq x$, we have

$$d_p(x_n,x) = d(x_n,p) + d(p,x) \geq d(p,x),$$

so that, for $d_p(x_n,x) \to 0$ to hold, $(x_n)_{n=1}^\infty$ must be eventually constant. Hence, for example, the Euclidean metric on $\mathbb{R}^n$ and—no matter how "Paris" is chosen—the corresponding French railroad metric are not equivalent. On the other hand, if (X,d) is discrete, then the identity from (X,d) to (X,d_p) is also continuous, so that d and d_p are equivalent.

(d) Let (X,d) be any metric space, and let $\tilde{d}$ be the metric defined in Example 2.1.2(f). We claim that d and $\tilde{d}$ are equivalent. The function

$$f\colon [0,\infty) \to [0,1), \quad t \mapsto \frac{t}{1+t}$$

is continuous and bijective with continuous inverse

$$g\colon [0,1) \to [0,\infty), \quad s \mapsto \frac{s}{1-s}.$$

Since $\tilde{d} = f \circ d$ (and, consequently, $d = g \circ \tilde{d}$), it follows that d and $\tilde{d}$ are indeed equivalent.

(e) Let (X,d) be any metric space, and let $U \subset X$ be open. Define

$$d_U(x,y) := d(x,y) + \left| \frac{1}{\operatorname{dist}(x, X \setminus U)} - \frac{1}{\operatorname{dist}(y, X \setminus U)} \right| \qquad (x,y \in U).$$

(If $U = X$, we formally set $\frac{1}{\operatorname{dist}(x,X\setminus U)} = \frac{1}{\operatorname{dist}(y,X\setminus U)} = \frac{1}{\infty} = 0$.) From Exercise 2.2.4, it follows that d_U is well defined on $U \times U$. We claim that d_U is a metric on U. Clearly, d_U is positive definite and symmetric. Let $x,y,z \in U$, and note that

$$\begin{aligned}
d_U(x,z) &= d(x,z) + \left| \frac{1}{\operatorname{dist}(x, X \setminus U)} - \frac{1}{\operatorname{dist}(z, X \setminus U)} \right| \\
&\leq d(x,y) + d(y,z) + \left| \frac{1}{\operatorname{dist}(x, X \setminus U)} - \frac{1}{\operatorname{dist}(y, X \setminus U)} \right. \\
&\qquad\qquad \left. + \frac{1}{\operatorname{dist}(y, X \setminus U)} - \frac{1}{\operatorname{dist}(z, X \setminus U)} \right| \\
&\leq d(x,y) + \left| \frac{1}{\operatorname{dist}(x, X \setminus U)} - \frac{1}{\operatorname{dist}(y, X \setminus U)} + \right| \\
&\quad + d(y,z) + \left| \frac{1}{\operatorname{dist}(y, X \setminus U)} - \frac{1}{\operatorname{dist}(z, X \setminus U)} \right| \\
&= d_U(x,y) + d_U(y,z).
\end{aligned}$$

We claim that d restricted to $U \times U$ and d_U are equivalent. Since

$$d(x,y) \le d_U(x,y) \qquad (x,y \in U),$$

the continuity of the identity from (U, d_U) to (U, d) is clear. To prove that the identity on U is also continuous in the converse direction, first note that nothing has to be shown if $U = X$. We may thus suppose without loss of generality that $U \subsetneq X$. Let $(x_n)_{n=1}^\infty$ be a sequence in U that converges to $x \in U$ with respect to d; that is, $d(x_n, x) \to 0$. By Exercise 3 below, this entails that $\operatorname{dist}(x_n, X \setminus U) \to \operatorname{dist}(x, X \setminus U)$ and thus

$$d_U(x_n, x) = d(x_n, x) + \left| \frac{1}{\operatorname{dist}(x_n, X \setminus U)} - \frac{1}{\operatorname{dist}(x, X \setminus U)} \right| \to 0.$$

Hence, $(x_n)_{n=1}^\infty$ converges to x as well with respect to d_U.

Exercises

1. Let $((X_k, d_k))_{k=1}^\infty$ be a sequence of metric spaces, and let $X := \prod_{k=1}^\infty X_k$ be equipped with the metric d from Example 2.1.2(g). Show that convergence in X is coordinatewise convergence: a sequence $\left(\left(x_1^{(n)}, x_2^{(n)}, x_3^{(n)}, \ldots\right)\right)_{n=1}^\infty$ in X converges to $(x_1, x_2, x_3, \ldots) \in X$ with respect to d if and only if $x_k^{(n)} \to x_k$ for each $k \in \mathbb{N}$.
2. Let (X, d_X) and (Y, d_Y) be metric spaces, let $p \in X$, and let d_p denote the corresponding French railroad metric on X. Show that $f\colon X \to Y$ is continuous with respect to d_p if and only if it is continuous at p with respect to d_X.
3. Let (X, d) be a metric space, and let $\varnothing \ne S \subset X$. Show that the function

$$X \to \mathbb{R}, \quad x \mapsto \operatorname{dist}(x, S)$$

is continuous.
4. Let E and F be normed spaces, and let $T\colon E \to F$ be linear. Show that the following are equivalent.
   (i) T is continuous;
   (ii) T is continuous at 0;
   (iii) There is $C \ge 0$ such that $\|T(x)\| \le C\|x\|$ for all $x \in E$.
5. Let E and F be normed spaces, let $T: E \to F$ be linear, and suppose that $\dim E < \infty$. Show that T is continuous. (*Hint*: For $x \in E$, define $|||x||| := \max\{\|x\|, \|T(x)\|\}$; show that $|||\cdot|||$ is a norm on E, and use Proposition B.1.)
6. On $C([0,1], \mathbb{F})$ we have the two norms $\|\cdot\|_1$ and $\|\cdot\|_\infty$ introduced in Example 2.1.2(c). Show that the metrics induced by these two norms are not equivalent.

2.4 Completeness

As we can define convergent sequences in metric spaces, we can speak of Cauchy sequences.

Definition 2.4.1. *Let (X, d) be a metric space. A sequence $(x_n)_{n=1}^\infty$ in X is called a* Cauchy sequence *if, for each $\epsilon > 0$, there is $n_\epsilon > 0$ such that $d(x_n, x_m) < \epsilon$ for all $n, m \geq n_\epsilon$.*

As in $\mathbb{R}^n$, we have the following.

Proposition 2.4.2. *Let (X, d) be a metric space, and let $(x_n)_{n=1}^\infty$ be a convergent sequence in X. Then $(x_n)_{n=1}^\infty$ is a Cauchy sequence.*

Proof. Let $x := \lim_{n\to\infty} x_n$, and let $\epsilon > 0$. Then there is $n_\epsilon > 0$ such that $d(x_n, x) < \frac{\epsilon}{2}$ for all $n \geq n_\epsilon$. Consequently, we have

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \qquad (n, m \geq n_\epsilon),$$

so that $(x_n)_{n=1}^\infty$ is a Cauchy sequence. □

In $\mathbb{R}^n$, the converse holds as well: Every Cauchy sequence converges. For general metric spaces, this is clearly false: the sequence $\left(\frac{1}{n}\right)_{n=1}^\infty$ is a Cauchy sequence in the metric space $(0, 1)$—equipped with its canonical metric—but has no limit in that space. This makes the following definition significant.

Definition 2.4.3. *A metric space (X, d) is called* complete *if every Cauchy sequence in X converges.*

A normed space that is complete with respect to the metric induced by its norm is also called a *Banach space*.

Examples 2.4.4. (a) $\mathbb{R}^n$ is complete.

(b) In a discrete metric space, every Cauchy sequence is eventually constant and therefore convergent. Hence, discrete metric spaces are complete.

(c) Let $S \neq \varnothing$ be a set, and let (Y, d) be a complete metric space. We claim that the metric space $(B(S, Y), D)$ from Example 2.1.2(d) is complete. Let $(f_n)_{n=1}^\infty$ be a Cauchy sequence in $B(S, Y)$. Let $\epsilon > 0$, and choose $n_\epsilon > 0$ such that $D(f_n, f_m) < \epsilon$ for all $n, m \geq n_\epsilon$. For $x \in S$, we then have

$$d(f_n(x), f_m(x)) \leq D(f_n, f_m) < \epsilon \qquad (n, m \geq n_\epsilon).$$

Consequently, $(f_n(x))_{n=1}^\infty$ is a Cauchy sequence in Y for each $x \in S$. Since Y is complete, we can therefore define $f : S \to Y$ by letting

$$f(x) := \lim_{n\to\infty} f_n(x) \qquad (x \in S).$$

We first claim that f lies in $B(S, Y)$ and is, in fact, the limit of $(f_n)_{n=1}^\infty$ with respect to D. To see this, let $x \in S$, and note that

$$d(f_n(x), f(x)) = \lim_{m\to\infty} d(f_n(x), f_m(x))$$

for $n \in \mathbb{N}$ by Example 2.3.9. It follows for $n \geq n_\epsilon$ that

$$
\begin{aligned}
&d(f_n(x), f(x)) \\
&= \lim_{m\to\infty} d(f_n(x), f_m(x)) \le \limsup_{m\to\infty} D(f_n, f_m) \le \epsilon \qquad (x \in S).
\end{aligned}
$$

Let $n \ge n_\epsilon$, and let $C := \sup_{x,y\in S} d(f_n(x), f_n(y))$, which is finite by the definition of $B(S, Y)$. From the previous inequality, we obtain, for arbitrary $x, y \in S$, that

$$d(f(x), f(y)) \le d(f(x), f_n(x)) + d(f_n(x), f_n(y)) + d(f_n(y), f(y)) \le 2\epsilon + C.$$

Hence, f belongs to $B(S, Y)$. Since $d(f_n(x), f(x)) \le \epsilon$ for all $x \in S$ and $n \ge n_\epsilon$, we eventually obtain:

$$D(f_n, f) = \sup_{x\in S} d(f_n(x), f(x)) \le \epsilon \qquad (n \ge n_\epsilon).$$

This is sufficient to guarantee that $f = \lim_{n\to\infty} f_n$ in $(B(S, Y), D)$.

The following proposition indicates how to get new complete spaces from old ones.

Proposition 2.4.5. *Let (X, d) be a metric space, and let Y be a subspace of X.*

(i) *If X is complete and if Y is closed in X, then Y is complete.*
(ii) *If Y is complete, then it is closed in X.*

Proof. Suppose that X is complete and that Y is closed in X. Let $(x_n)_{n=1}^\infty$ be a Cauchy sequence in Y. Then $(x_n)_{n=1}^\infty$ is also a Cauchy sequence in X and thus has a limit $x \in X$. Since Y is closed, Corollary 2.3.5 yields that $x \in Y$, so that Y is complete. This proves (i).

For (ii), let $(y_n)_{n=1}^\infty$ be a sequence in Y that converges to $y \in X$. Since $(y_n)_{n=1}^\infty$ converges in X, it is a Cauchy sequence in X and thus in Y. Since Y is complete, there is $y' \in Y$ with $y' = \lim_{n\to\infty} y_n$. If $(y_n)_{n=1}^\infty$ converges to y' in Y, it does so in X. Uniqueness of the limit yields that $y' = y$. Hence, y lies in Y. Corollary 2.3.5 thus yields that Y is closed in X. □

Example 2.4.6. Let (X, d_X) and (Y, d_Y) be metric spaces. We define

$$C(X, Y) := \{f : X \to Y : f \text{ is continuous}\}$$

and

$$C_b(X, Y) := B(X, Y) \cap C(X, Y).$$

Clearly, $C_b(X, Y)$ is a subspace of the metric space $(B(X, Y), D)$. We claim that $C_b(X, Y)$ is closed in $B(X, Y)$ and therefore complete if (Y, d_Y) is. Let $(f_n)_{n=1}^\infty$ be a sequence in $C_b(X, Y)$ that converges to $f \in B(X, Y)$. We claim that f is again continuous. To see this, fix $x_0 \in X$. We show that f is continuous at x_0. Let $\epsilon > 0$. Since $f_n \to f$ in $B(X, Y)$, there is $n_\epsilon \in \mathbb{N}$ such that

$D(f_n, f) < \frac{\epsilon}{3}$ for $n \geq n_\epsilon$. Fix $n \geq n_\epsilon$. Since f_n is continuous at x_0, the set $N := f_n^{-1}(B_{\frac{\epsilon}{3}}(f_n(x_0)))$ is a neighborhood of x_0. Let $x \in N$, and note that

$$\begin{aligned} d(f(x), f(x_0)) &\leq d(f(x), f_n(x)) + d(f_n(x), f_n(x_0)) + d(f_n(x_0), f(x_0)) \\ &\leq D(f_n, f) + d(f_n(x), f_n(x_0)) + D(f_n, f) \\ &< \frac{2\epsilon}{3} + d(f_n(x), f_n(x_0)), \qquad \text{because } n \geq n_\epsilon, \\ &< \epsilon, \qquad \text{because } x \in N. \end{aligned}$$

It follows that $N \subset f^{-1}(B_\epsilon(f(x_0)))$, so that $f^{-1}(B_\epsilon(f(x_0))) \in \mathcal{N}_{x_0}$. Since $\epsilon > 0$ was arbitrary, this is enough to guarantee the continuity of f at x_0.

In view of Proposition 2.4.5, the following assertion seems to defy reason at first glance.

Proposition 2.4.7. *Let (X, d) be a complete metric space, and let $U \subset X$ be open. Then (U, d_U) is a complete metric space, where d_U is defined as in Example* 2.3.13(e).

Proof. If $U = X$, we have $d_U = d$, so that the claim is trivially true. Hence, suppose that $U \subsetneq X$.

Let $(x_n)_{n=1}^\infty$ be a Cauchy sequence in (U, d_U). Then $(x_n)_{n=1}^\infty$ is easily seen to be a Cauchy sequence in (X, d) as well. Let $x \in X$ be its limit in (X, d). We first claim that $x \in U$. Assume towards a contradiction that $x \in X \setminus U$. From Exercise 2.3.3, we conclude that $\operatorname{dist}(x_n, X \setminus U) \to 0$. Since $(x_n)_{n=1}^\infty$ is a Cauchy sequence in (U, d_U), there is $n_1 \in \mathbb{N}$ such that

$$\left| \frac{1}{\operatorname{dist}(x_n, X \setminus U)} - \frac{1}{\operatorname{dist}(x_m, X \setminus U)} \right| \leq d_U(x_n, x_m) \leq 1 \qquad (n, m \geq n_1).$$

Fix $m \geq n_1$, and note that therefore

$$\begin{aligned} \frac{1}{\operatorname{dist}(x_n, X \setminus U)} &\leq \left| \frac{1}{\operatorname{dist}(x_n, X \setminus U)} - \frac{1}{\operatorname{dist}(x_m, X \setminus U)} \right| + \frac{1}{\operatorname{dist}(x_m, X \setminus U)} \\ &\leq 1 + \frac{1}{\operatorname{dist}(x_m, X \setminus U)} \qquad (n \geq n_1). \end{aligned}$$

This is impossible, however, if $\operatorname{dist}(x_n, X \setminus U) \to 0$. Consequently, $x \in U$ must hold.

Since d and d_U are equivalent on U, we see that $d_U(x_n, x) \to 0$ as well. Hence, x is the limit of $(x_n)_{n=1}^\infty$ in (U, d_U). $\square$

At first glance, Proposition 2.4.7 seems to be paradoxical, to say the least. Any open subset of a complete metric space is supposed to be complete with respect to an equivalent metric. Doesn't this and Proposition 2.4.5(ii) immediately yield that every open subset of a complete metric space is also closed?

This is clearly wrong. The apparent paradox is resolved if one recalls the definition of a sub*space* of a metric space: (U, d_U) is not a subspace of the metric space (X, d), even though the two metrics d and d_U are equivalent on U.

We now present a famous property of complete metric spaces, for which we first require a definition.

Definition 2.4.8. *Let* (X, d) *be a metric space. The* diameter *of a subset* $S \neq \varnothing$ *of* X *is defined as*

$$\operatorname{diam}(S) := \sup\{d(x, y) : x, y \in S\}.$$

Theorem 2.4.9 (Cantor's intersection theorem). *Let* (X, d) *be a complete metric space, and let* $(F_n)_{n=1}^{\infty}$ *be a sequence of nonempty closed subsets of* X *such that* $F_1 \supset F_2 \supset F_3 \supset \cdots$ *and* $\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$*. Then* $\bigcap_{n=1}^{\infty} F_n$ *contains precisely one point of* X*.*

Proof. For each $n \in \mathbb{N}$, let $x_n \in F_n$. We claim that the sequence $(x_n)_{n=1}^{\infty}$ is a Cauchy sequence. To see this, let $\epsilon > 0$. Choose $n_\epsilon \in \mathbb{N}$ such that $\operatorname{diam}(F_n) < \epsilon$ for $n \geq n_\epsilon$. Let $n, m \geq n_\epsilon$. Since the sequence $(F_n)_{n=1}^{\infty}$ is decreasing, it follows that $x_n, x_m \in F_{n_\epsilon}$, so that

$$d(x_n, x_m) \leq \operatorname{diam}(F_{n_\epsilon}) < \epsilon.$$

Consequently, $(x_n)_{n=1}^{\infty}$ is indeed a Cauchy sequence and therefore converges in X, to x say. Since $x_m \in F_m \subset F_n$ for all $n, m \in \mathbb{N}$, with $m \geq n$, it follows from Corollary 2.3.5 that $x = \lim_{m\to\infty} x_m \in F_n$ for all $n \in \mathbb{N}$ and thus $x \in \bigcap_{n=1}^{\infty} F_n$.

To show that $\bigcap_{n=1}^{\infty} F_n = \{x\}$, assume towards a contradiction that there is $x' \in \bigcap_{n=1}^{\infty} F_n$ different from x. Let $\epsilon_0 := d(x, x') > 0$, and choose $n \in \mathbb{N}$ so large that $\operatorname{diam}(F_n) < \epsilon_0$. Since $x, x' \in F_n$, we obtain

$$d(x, x') \leq \operatorname{diam}(F_n) < \epsilon_0 = d(x, x'),$$

which is impossible. $\square$

Next, we show that *any* metric space is—in a sense yet to be made precise—already a subspace of a complete metric space.

Definition 2.4.10. *Let* (X, d) *be a metric space. A* completion *of* (X, d) *is a metric space* $\left(\tilde{X}, \tilde{d}\right)$ *together with a map* $\iota : X \to \tilde{X}$ *with the following properties.*

(a) $\left(\tilde{X}, \tilde{d}\right)$ *is complete;*
(b) $\tilde{d}(\iota(x), \iota(y)) = d(x, y)$ *for* $x, y \in X$*;*
(c) $\iota(X)$ *is dense in* $\tilde{X}$*.*

We show that, first of all, *every* metric space has a completion and, secondly, that this completion is unique (in a certain sense).

To specify what we mean by uniqueness of a completion, we require another definition.

Definition 2.4.11. *Let* (X, d_X) *and* (X, d_Y) *be metric spaces. A function* $f\colon X \to Y$ *is called an* isometry *(or* isometric*) if*

$$d_Y(f(x), f(y)) = d_X(x, y) \qquad (x, y \in X).$$

If f *is also bijective, we call* f *an* isometric isomorphism.

Lemma 2.4.12. *Let* (X, d) *be a metric space, let* $\left(\tilde{X}_1, \tilde{d}_1\right)$ *and* $\left(\tilde{X}_2, \tilde{d}_2\right)$ *be completions of* (X, d)*, and let* $\iota_1 : X \to \tilde{X}_1$ *and* $\iota_2 : X \to \tilde{X}_2$ *denote the corresponding maps from Definition* 2.4.10. *Then there is a unique isometric isomorphism* $f\colon \tilde{X}_1 \to \tilde{X}_2$ *such that* $f \circ \iota_1 = \iota_2$.

Proof. We begin with the definition of f. Let $x \in \tilde{X}_1$. Since $\iota_1(X)$ is dense in $\tilde{X}_1$, there is a sequence $(x_n)_{n=1}^\infty$ in X such that $x = \lim_{n\to\infty} \iota_1(x_n)$. It is clear that $(\iota_1(x_n))_{n=1}^\infty$ is a Cauchy sequence in $\tilde{X}_1$, and Definition 2.4.10(b) implies that $(x_n)_{n=1}^\infty$ is a Cauchy sequence in X. Again Definition 2.4.10(b) guarantees that $(\iota_2(x_n))_{n=1}^\infty$ is a Cauchy sequence in $\tilde{X}_2$ and therefore converges. Let $f(x) := \lim_{n\to\infty} \iota_2(x_n)$.

We first prove that f is *well defined*, that is, does not depend on the particular choice of a sequence $(x_n)_{n=1}^\infty$. To prove this, let $(x_n')_{n=1}^\infty$ be another sequence in X with $x = \lim_{n\to\infty} \iota_1(x_n')$. It follows that

$$d(x_n, x_n') = \tilde{d}_1(\iota_1(x_n), \iota_1(x_n')) \le \tilde{d}_1(\iota_1(x_n), x) + \tilde{d}_1(x, \iota_1(x_n')) \to 0$$

and therefore

$$\begin{aligned} \tilde{d}_2(\iota_2(x_n'), f(x)) &\le \tilde{d}_2(\iota_2(x_n'), \iota_2(x_n)) + \tilde{d}_2(\iota_2(x_n), f(x)) \\ &= d(x_n', x_n) + \tilde{d}_2(\iota_2(x_n), f(x)) \\ &\to 0. \end{aligned}$$

All in all, $f(x) = \lim_{n\to\infty} \iota_2(x_n')$ holds, so that f is indeed well defined.

Next, we prove that f is an isometry. Let $x, y \in \tilde{X}_1$ and let $(x_n)_{n=1}^\infty$ and $(y_n)_{n=1}^\infty$ be the corresponding sequences in X used to define $f(x)$ and $f(y)$, respectively. From

$$\begin{aligned} \tilde{d}_2(f(x), f(y)) &= \lim_{n\to\infty} \tilde{d}_2(\iota_2(x_n), \iota_2(x_n)) \\ &= \lim_{n\to\infty} d(x_n, y_n) \\ &= \lim_{n\to\infty} \tilde{d}_1(\iota_1(x_n), \iota_1(x_n)) \\ &= \tilde{d}_1(x, y), \end{aligned}$$

we see that f is isometric. This immediately also proves the injectivity of f.

Clearly, $f \circ \iota_1 = \iota_2$ holds, so that $f\left(\tilde{X}_1\right) \supset \iota_2(X)$ must be dense in $\tilde{X}_2$. We claim that $f\left(\tilde{X}_1\right)$ is a complete subspace of $\tilde{X}_2$ and therefore closed (this

implies that $f\left(\tilde{X}_1\right)$ must be all of $\tilde{X}_2$). Let $(x_n)_{n=1}^\infty$ be a sequence in $\tilde{X}_1$ such that $(f(x_n))_{n=1}^\infty$ is a Cauchy sequence in $\tilde{X}_2$. Since f is an isometry, $(x_n)_{n=1}^\infty$ is also a Cauchy sequence in $\tilde{X}_1$ and thus convergent to some $x \in \tilde{X}_1$. Again since f is an isometry, it follows that $\lim_{n\to\infty} f(x_n) = f(x)$ in $\tilde{X}_2$.

Finally, to prove the uniqueness of f, let $\tilde{f}\colon \tilde{X}_1 \to \tilde{X}_2$ be another map as described in the statement of the lemma. Let $x \in \tilde{X}_1$. By Definition 2.4.10(c), there is a sequence $(x_n)_{n=1}^\infty$ in X with $\lim_{n\to\infty} \iota_1(x_n) = x$. We obtain that

$$f(x) = \lim_{n\to\infty} f(\iota_1(x_n)) = \lim_{n\to\infty} \iota_2(x_n) = \lim_{n\to\infty} \tilde{f}(\iota_1(x_n)) = \tilde{f}(x).$$

Since $x \in \tilde{X}_1$ was arbitrary, this proves that $f = \tilde{f}$. □

In less formal (but probably much more digestible) language, Lemma 2.4.12 asserts that a completion of a metric space (if it exists at all!) is unique up to isometric isomorphism.

The existence of the completion of a given metric space is surprisingly easy to establish.

Theorem 2.4.13. *Let (X,d) be a metric space. Then (X,d) has a completion, which is unique up to isometric isomorphism.*

Proof. In view of Lemma 2.4.12, only the existence of the completion still has to be shown. It is sufficient to find some complete metric space and an isometry ι from X into that space: just let $\tilde{X} := \overline{\iota(X)}$. The complete metric space into which we embed X is the Banach space $C_b(X,\mathbb{R})$.

Fix $x_0 \in X$. For $x \in X$, define

$$f_x\colon X \to \mathbb{R}, \quad t \mapsto d(x,t) - d(x_0,t).$$

In view of Example 2.3.9, it is clear that f_x is continuous for each $x \in X$, and also, due to the inequality $(***)$ from Example 2.3.9, we have

$$|f_x(t)| \leq d(x,x_0) + d(t,t) = d(x,x_0) \qquad (t \in X),$$

so that f_x lies even in $C_b(X,\mathbb{R})$. We claim that the map

$$\iota\colon X \to C_b(X,\mathbb{R}), \quad x \mapsto f_x$$

is an isometry. To see this, fix $x,y \in X$ and note that, by $(***)$ again,

$$D(\iota(x),\iota(y)) = \sup_{t\in X} |f_x(t) - f_y(t)| = \sup_{t\in X} |d(x,t) - d(y,t)| \leq d(x,y),$$

holds; on the other hand, we have

$$D(\iota(x),\iota(y)) = \sup_{t\in X} |f_x(t) - f_y(t)| \geq |f_x(y) - f_y(y)| = d(x,y),$$

which proves the claim. □

In view of the uniqueness of a completion up to isometric isomorphism, we are justified to speak of *the* completion of a metric space. For the sake of notational convenience, we also identify a metric space with its canonical image in its completion.

We now turn to one of the most fundamental theorems on complete metric spaces.

Theorem 2.4.14 (Bourbaki's Mittag-Leffler theorem). *Suppose that* $((X_n, d_n))_{n=0}^\infty$ *is a sequence of complete metric spaces, and let* $f_n \colon X_n \to X_{n-1}$ *for* $n \in \mathbb{N}$ *be continuous with dense range. Then*

$$\bigcap_{n=1}^\infty (f_1 \circ f_2 \circ \cdots \circ f_n)(X_n)$$

is dense in X_0.

Proof. We first inductively define new metrics $\tilde{d}_0, \tilde{d}_1, \tilde{d}_2, \ldots$ on the spaces $X_0, X_1, X_2, \ldots$ such that

- $\tilde{d}_n$ and d_n are equivalent for $n \in \mathbb{N}_0$,
- $\left(X_n, \tilde{d}_n\right)$ is complete for each $n \in \mathbb{N}_0$, and
- $\tilde{d}_{n-1}(f_n(x), f_n(y)) \leq \tilde{d}_n(x, y)$ for $n \in \mathbb{N}$ and $x, y \in X_n$.

This is accomplished by letting $\tilde{d}_0 := d_0$ and, once $\tilde{d}_0, \ldots, \tilde{d}_{n-1}$ have been defined for some $n \in \mathbb{N}$, letting

$$\tilde{d}_n(x, y) := d_n(x, y) + \tilde{d}_{n-1}(f_n(x), f_n(y)) \qquad (x, y \in X_n).$$

In what follows, we consider the spaces $X_0, X_1, X_2, \ldots$ equipped with the metrics $\tilde{d}_0, \tilde{d}_1, \tilde{d}_2, \ldots$ instead of with $d_0, d_1, d_2, \ldots$.

Let $U_0 \subset X$ be open and not empty. We need to show that

$$U_0 \cap \bigcap_{n=1}^\infty (f_1 \circ \cdots \circ f_n)(X_n) \neq \varnothing.$$

Since $f_1(X_1)$ is dense in X_0, there is $x_1 \in X_1$ with $f_1(x_1) \in U_0$. Since f_1 is continuous at x_1, there is $\delta_1 \in (0, 1]$ such that $\overline{f_1(B_{\delta_1}(x_1))} \subset U_0$. Let $U_1 := B_{\delta_1}(x_1)$. Since $f_2(X_2)$ is dense in X_1, there is $x_2 \in X_2$ with $f_2(x_2) \in U_1$. Since f_2 is continuous at x_2, there is $\delta_2 \in \left(0, \frac{1}{2}\right]$ such that $\overline{f_2(B_{\delta_2}(x_2))} \subset U_1$. Let $U_2 := B_{\delta_2}(x_2)$, and continue in this fashion.

We thus obtain a sequence $(U_n)_{n=1}^\infty$ of open balls such that $\overline{f_n(U_n)} \subset U_{n-1}$ for $n \in \mathbb{N}$ and such that U_n has radius at most $\frac{1}{n}$. For $n \in \mathbb{N}_0$ and $m \in \mathbb{N}$, let

$$Y_{n,m} := \overline{(f_{n+1} \circ \cdots \circ f_{n+m})(U_{n+m})}.$$

It follows that $Y_{n,m} \neq \varnothing$, that $\operatorname{diam}(Y_{n,m}) \leq \frac{2}{n+m}$, and that $Y_{n,m+1} \subset Y_{n,m}$. From Cantor's intersection theorem, it follows that there is $y_n \in \bigcap_{m=1}^\infty Y_{n,m}$.

From the construction, it is immediate that $f_n(y_n) = y_{n-1}$ and thus that $(f_1 \circ \cdots \circ f_n)(y_n) = y_0$ for $n \in \mathbb{N}$. Consequently,

$$y_0 \in U_0 \cap \bigcap_{n=1}^{\infty} (f_1 \circ \cdots \circ f_n)(X_n)$$

holds. □

The name "Mittag-Leffler theorem" for Theorem 2.4.14 may sound bewildering, but the well-known Mittag-Leffler theorem from complex analysis (Theorem A.1) can be obtained as a consequence of it (see Appendix A; besides some background from complex variables, you will also need material from Sections 3.1 to 3.4 for it). We turn, however, to another consequence of Theorem 2.4.14.

Lemma 2.4.15. *Let (X, d) be a metric space, and let $U_1, \ldots, U_n \subset X$ be dense open subsets of X. Then $U_1 \cap \cdots \cap U_n$ is dense in X.*

Proof. By induction, it is clear that we may limit ourselves to the case where $n = 2$. Let $x \in X$, and let $\epsilon > 0$. Since U_1 is dense in X, we have $B_\epsilon(x) \cap U_1 \neq \varnothing$. Since $B_\epsilon(x) \cap U_1$ is open—and thus a neighborhood of each of its points—it follows from the denseness of U_2 that $B_\epsilon(x) \cap U_1 \cap U_2 \neq \varnothing$. Since $\epsilon > 0$ was arbitrary, we conclude that $x \in \overline{U_1 \cap U_2}$. □

Theorem 2.4.16 (Baire's theorem). *Let (X, d) be a complete metric space, and let $(U_n)_{n=1}^{\infty}$ be a sequence of dense open subsets of X. Then $\bigcap_{n=1}^{\infty} U_n$ is dense in X.*

Proof. By Lemma 2.4.15, we may replace U_n by $U_1 \cap \cdots \cap U_n$ and thus suppose without loss of generality that $U_1 \supset U_2 \supset \cdots$. Let $(X_0, d_0) := (X, d)$, and let $(X_n, d_n) := (U_n, d_{U_n})$, where d_{U_n} is defined for $n \in \mathbb{N}$ as in Example 2.3.13(e). Furthermore, let $f_n : X_n \to X_{n-1}$ be the inclusion map for $n \in \mathbb{N}$. Since d and d_{U_n} are equivalent on X_n for $n \in \mathbb{N}$, it is clear that $f_1, f_2, \ldots$ are continuous. By the hypothesis, (X_0, d_0) is complete and the same is true for (X_n, d_n) with $n \in \mathbb{N}$ by Proposition 2.4.7. It follows from Theorem 2.4.14 that

$$\bigcap_{n=1}^{\infty} (f_1 \circ \cdots \circ f_n)(X_n) = \bigcap_{n=1}^{\infty} U_n$$

is dense in $X_0 = X$. □

The following is an immediate consequence of Baire's theorem (just pass to complements).

Corollary 2.4.17. *Let (X, d) be a complete metric space, and let $(F_n)_{n=1}^{\infty}$ be a sequence of closed subsets of X such that $\bigcup_{n=1}^{\infty} F_n$ has a nonempty interior. Then at least one of the sets $F_1, F_2, \ldots$ has a nonempty interior.*

To illustrate the power of Baire's theorem, we turn to an example from elementary calculus. We all know that there are continuous functions that are not differentiable at certain points (take the absolute value function, for instance), and it is not very hard to come up with continuous functions that are not differentiable at a finite, and even countable, number of points. But is there a continuous function, on an interval say, that fails to be differentiable at each point of its domain? The following example gives the answer.

Example 2.4.18. For $n \in \mathbb{N}$, let F_n consist of those $f \in C([0,2],\mathbb{R})$ for which there is $t \in [0,1]$ such that

$$\sup_{h\in(0,1)} \frac{|f(t+h)-f(t)|}{h} \leq n.$$

Obviously, if $f \in C([0,2],\mathbb{R})$ is differentiable at some point $t \in [0,1]$, then

$$\sup_{h\in(0,1)} \frac{|f(t+h)-f(t)|}{h} < \infty$$

must hold, so that $f \in \bigcup_{n=1}^{\infty} F_n$. Hence, if every continuous function on $[0,2]$ is differentiable at some point of $[0,1]$, we have $C([0,2],\mathbb{R}) = \bigcup_{n=1}^{\infty} F_n$. Using Corollary 2.4.17, we show that this is not possible.

To be able to apply Corollary 2.4.17, we first need to show that the sets F_n for $n \in \mathbb{N}$ are closed in $C([0,2],\mathbb{R})$. Fix $n \in \mathbb{N}$, and let $(f_m)_{m=1}^{\infty}$ be a sequence in F_n such that $\|f_m - f\|_\infty \to 0$ for some $f \in C([0,2],\mathbb{R})$. For each $m \in \mathbb{N}$, there is $t_m \in [0,1]$ such that

$$\sup_{h\in(0,1)} \frac{|f_m(t_m+h)-f_m(t_m)|}{h} \leq n.$$

Suppose without loss of generality that $(t_m)_{m=1}^{\infty}$ converges to some $t \in [0,1]$ (otherwise, replace $(t_m)_{m=1}^{\infty}$ by a convergent subsequence). Fix $h \in (0,1)$ and $\epsilon > 0$, and choose $m_\epsilon \in \mathbb{N}$ so large that

$$\left.\begin{Bmatrix} |f(t+h)-f(t_m+h)| \\ \|f-f_m\|_\infty \\ |f(t_m)-f(t)| \end{Bmatrix}\right. < \frac{\epsilon}{4}h \qquad (m \geq m_\epsilon).$$

For $m \geq m_\epsilon$, this implies

$$\begin{aligned} &|f(t+h)-f(t)| \\ &\leq \underbrace{|f(t+h)-f(t_m+h)|}_{<\frac{\epsilon}{4}h} + \underbrace{|f(t_m+h)-f_m(t_m+h)|}_{<\frac{\epsilon}{4}h} \\ &\quad + \underbrace{|f_m(t_m+h)-f_m(t_m)|}_{\leq nh} + \underbrace{|f_m(t_m)-f(t_m)|}_{<\frac{\epsilon}{4}h} + \underbrace{|f(t_m)-f(t)|}_{<\frac{\epsilon}{4}h} \\ &\leq nh + \epsilon h, \end{aligned}$$

so that

$$\frac{|f(t+h)-f(t)|}{h} \leq n+\epsilon.$$

Since h and ϵ were arbitrary, this means that $f \in F_n$. Hence, F_n is closed.

Assume towards a contradiction that every $f \in C([0,2],\mathbb{R})$ is differentiable at some point in $[0,1]$, so that $C([0,2],\mathbb{R}) = \bigcup_{n=1}^{\infty} F_n$. By Corollary 2.4.17, there are $n_0 \in \mathbb{N}$, $f \in C([0,2],\mathbb{R})$, and $\epsilon > 0$ such that $B_\epsilon(f) \subset F_{n_0}$. By the Weierstraß approximation theorem (Corollary 4.3.8 below), $B_\epsilon(f)$ contains at least one polynomial, say p. Since $B_\epsilon(f)$ is open, there is $\delta > 0$ such that $B_\delta(p) \subset B_\epsilon(f) \subset F_{n_0}$. Replacing f by p and ϵ by δ, we can thus suppose without loss of generality that f is continuously differentiable on $[0,2]$.

For $k \in \mathbb{N}$ and $j = 0,\ldots,k$, let $t_j := \frac{2j}{k}$. Define a "sawtooth function" $g_k \colon [0,2] \to \mathbb{R}$ by letting

$$g_k(t) := \begin{cases} \frac{\epsilon}{2}k(t-t_{j-1}), & t \in \left[t_{j-1}, t_{j-1}+\frac{1}{k}\right], \\ \frac{\epsilon}{2}k(t_j-t), & t \in \left[t_j-\frac{1}{k}, t_j\right] \end{cases}$$

for $j = 1,\ldots,n$ and $t \in [t_{j-1}, t_j]$.

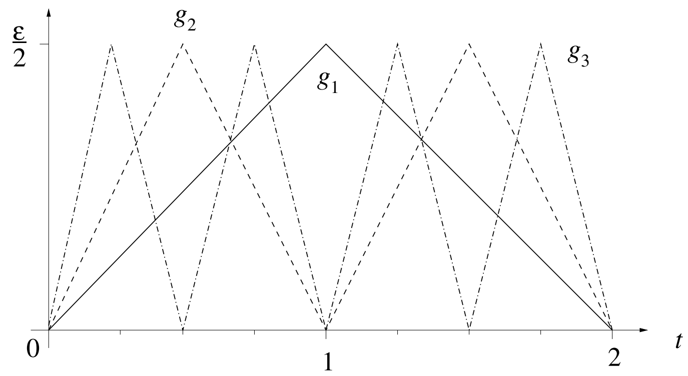

Fig. 2.3: Sawtooth functions

Then g_k is continuous with $\|g_k\|_\infty = \frac{\epsilon}{2}$, but

$$\sup_{h \in (0,1)} \frac{|g_k(t+h)-g_k(t)|}{h} = \frac{\epsilon}{2}k \tag{†}$$

holds for any $t \in [0,1]$. Since $f + g_k \in B_\epsilon(f) \subset F_{n_0}$, there is $t \in [0,1]$ such that

$$\sup_{h \in (0,1)} \frac{|(f+g_k)(t+h)-(f+g_k)(t)|}{h} \leq n_0.$$

This, however, yields

$$\begin{aligned}
&\sup_{h\in(0,1)} \frac{|g_k(t+h)-g_k(t)|}{h} \\
&\le \sup_{h\in(0,1)} \frac{|(f+g_k)(t+h)-(f+g_k)(t)|}{h} + \sup_{h\in(0,1)} \frac{|f(t+h)-f(t)|}{h} \\
&= n_0 + \|f'\|_\infty,
\end{aligned}$$

which contradicts (†) if we choose $k \in \mathbb{N}$ so large that $\frac{\epsilon}{2}k > n_0 + \|f'\|_\infty$.

Hence, the sets $F_1, F_2, \ldots$ have an empty interior, thus their union $\bigcup_{n=1}^\infty F_n$ cannot be all of $C([0,2],\mathbb{R})$, and consequently there must be a continuous function on $[0,1]$ that is nowhere differentiable.

Exercises

1. Let (X,d) be any metric space, let $p \in X$, and let d_p be the corresponding French railroad metric. Show that (X,d_p) is complete.
2. Let (X,d) be a complete metric space, and let $(x_n)_{n=0}^\infty$ be a sequence in X such that there is $\theta \in (0,1)$ with $d(x_{n+1},x_n) \le \theta\, d(x_n,x_{n-1})$ for $n \in \mathbb{N}$. Show that $(x_n)_{n=0}^\infty$ is convergent.
3. Use the previous problem to prove *Banach's fixed point theorem*: if (X,d) is a complete metric space, and if $f\colon X \to X$ is such that
$$d(f(x),f(y)) \le \theta\, d(x,y) \qquad (x,y \in X)$$
for some $\theta \in (0,1)$, then there is a unique $x \in X$ with $f(x) = x$.
4. Let (X,d) be a metric space, and let $\varnothing \neq S \subset X$. Show that
$$\operatorname{diam}(S) = \inf\{r > 0 : S \subset B_r(x) \text{ for all } x \in S\}.$$
5. Give an example showing that the demand that $\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$ in Cantor's intersection theorem cannot be dropped if we still want $\bigcap_{n=1}^\infty F_n \neq \varnothing$ to hold.
6. Let E be a normed space with a countable Hamel basis. Show that E is a Banach space if and only if $\dim E < \infty$. (*Hint*: You may use the fact that all finite-dimensional subspaces of a normed space are closed (Corollary B.3); then use Corollary 2.4.17.)
7. Let $(f_k)_{k=1}^\infty$ be a sequence in $C([0,1],\mathbb{F})$ that converges *pointwise* to a function $f\colon [0,1] \to \mathbb{F}$.
   (a) For $\theta > 0$ and $n \in \mathbb{N}$, let
   $$F_n := \{t \in [0,1] : |f_n(t) - f_k(t)| \le \theta \text{ for all } k \ge n\}.$$
   Show that F_n is closed, and that $[0,1] = \bigcup_{n=1}^\infty F_n$.
   (b) Let $\epsilon > 0$, and let I be a nondegenerate, closed subinterval of $[0,1]$. Show that there is a nondegenerate, closed interval J contained in $\mathring{I}$ such that
   $$|f(t) - f(s)| \le \epsilon \qquad (t,s \in J).$$
   (*Hint*: Apply (a) with $\theta := \frac{\epsilon}{3}$ and Corollary 2.4.17.)

(c) Let I be a nondegenerate closed subinterval of $[0,1]$. Show that there is a sequence $(I_n)_{n=1}^\infty$ of nondegenerate closed subintervals of I with $I_1 \supset \overset{\circ}{I}_1 \supset I_2 \supset \overset{\circ}{I}_2 \supset I_3 \supset \cdots$ such that
- The length of I_n is at most $\frac{1}{n}$, and
- $|f(t) - f(s)| \leq \frac{1}{n}$ for all $s, t \in I_n$.

What can be said about f at all points in $\bigcap_{n=1}^\infty I_n$?

(d) Conclude that the set of points in $[0,1]$ at which f is continuous is dense in $[0,1]$.

2.5 Compactness for Metric Spaces

The notion of compactness is one of the most crucial in all of topology (and one of the hardest to grasp).

Definition 2.5.1. *Let (X,d) be a metric space, and let $S \subset X$. An* open cover *for S is a collection $\mathcal{U}$ of open subsets of X such that $S \subset \bigcup\{U : U \in \mathcal{U}\}$.*

Definition 2.5.2. *A subset K of a metric space (X,d) is called* compact *if, for each open cover $\mathcal{U}$ of K, there are $U_1, \ldots, U_n \in \mathcal{U}$ such that $K \subset U_1 \cup \cdots \cup U_n$.*

Definition 2.5.2 is often worded as, "A set is compact if and only if each open cover has a finite subcover."

Examples 2.5.3. (a) Let (X,d) be a metric space, and let $S \subset X$ be finite; that is, $S = \{x_1, \ldots, x_n\}$. Let $\mathcal{U}$ be an open cover of X. Then, for each $j = 1, \ldots, n$, there is $U_j \in \mathcal{U}$ such that $x_j \in U_j$. It follows that $S \subset U_1 \cup \cdots \cup U_n$. Hence, S is compact.

(b) Let (X,d) be a compact metric space, and let $\varnothing \neq K \subset X$ be compact. Fix $x_0 \in K$. Since $\{B_r(x_0) : r > 0\}$ is an open cover of K, there are $r_1, \ldots, r_n > 0$ such that

$$K \subset B_{r_1}(x_0) \cup \cdots \cup B_{r_n}(x_0).$$

With $R := \max\{r_1, \ldots, r_n\}$, we see that $K \subset B_R(x_0)$, so that $\operatorname{diam}(K) \leq 2R < \infty$. This means, for example, that any unbounded subset of $\mathbb{R}^n$ (or, more generally, of any normed space) cannot be compact. In particular, the only compact normed space is $\{0\}$.

(c) Let $X = (0,1)$ be equipped with the usual metric. For $r \in (0,1)$, let $U_r := (r,1)$. Then $\{U_r : r \in (0,1)\}$ is an open cover for $(0,1)$ which has no finite subcover.

Before we turn to more (and more interesting) examples of compact metric spaces, we establish a few hereditary properties.

Proposition 2.5.4. *Let (X,d) be a metric space, and let Y be a subspace of X.*

(i) *If X is compact and Y is closed in X, then Y is compact.*
(ii) *If Y is compact, then it is closed in X.*

Proof. For (i), let $\mathcal{U}$ be an open cover for Y. Since Y is closed in X, the family $\mathcal{U} \cup \{X \setminus Y\}$ is an open cover for X. Since X is compact, it has a finite subcover, i.e., there are $U_1, \dots, U_n \in \mathcal{U}$ such that

$$X = U_1 \cup \cdots \cup U_n \cup X \setminus Y.$$

Taking the intersection with Y, we see that $Y \subset U_1 \cup \cdots \cup U_n$.

For (ii), let $x \in X \setminus Y$. For each $y \in Y$, there are $\epsilon_y, \delta_y > 0$ such that $B_{\epsilon_y}(x) \cap B_{\delta_y}(y) = \varnothing$. Since $\{B_{\delta_y}(y) : y \in Y\}$ is an open cover for Y, there are $y_1, \dots, y_1 \in Y$ such that

$$Y \subset B_{\delta_{y_1}}(y_1) \cup \cdots \cup B_{\delta_{y_n}}(y_n).$$

Letting $\epsilon := \min\{\epsilon_{y_1}, \dots, \epsilon_{y_n}\}$, we obtain that

$$B_\epsilon(x) \cap Y \subset B_\epsilon(x) \cap \left(B_{\delta_{y_1}}(y_1) \cup \cdots \cup B_{\delta_{y_n}}(y_n)\right) = \varnothing$$

and thus $B_\epsilon(x) \subset X \setminus Y$. Since $x \in X \setminus Y$ was arbitrary, this means that $X \setminus Y$ is open. □

Proposition 2.5.5. *Let (K, d_K) be a compact metric space, let (Y, d_Y) be any metric space, and let $f : K \to Y$ be continuous. Then $f(K)$ is compact.*

Proof. Let $\mathcal{U}$ be an open cover for $f(K)$. Then $\{f^{-1}(U) : U \in \mathcal{U}\}$ is an open cover for K by Corollary 2.3.10. Hence, there are $U_1, \dots, U_n \in \mathcal{U}$ with

$$K = f^{-1}(U_1) \cup \cdots \cup f^{-1}(U_n)$$

and thus

$$f(K) \subset U_1 \cup \cdots \cup U_n.$$

This proves the claim. □

Corollary 2.5.6. *Let (K, d) be a non-empty, compact metric space, and let $f : K \to \mathbb{R}$ be continuous. Then f attains both a minimum and a maximum on K.*

Proof. Let $M := \sup f(K)$. Since $f(K)$ is compact, it is bounded, so that $M < \infty$. For each $n \in \mathbb{N}$, there is $y_n \in f(K)$ such that $y_n > M - \frac{1}{n}$; it is clear that $M = \lim_{n\to\infty} y_n$. Since $f(K)$ is closed in $\mathbb{R}$, it follows that $M \in f(K)$. Hence, there is $x_0 \in K$ such that $f(x_0) = M$.

An analogous argument works for $\inf f(K)$. □

The real line $\mathbb{R}$ has the Bolzano–Weierstraß property: every bounded sequence in $\mathbb{R}$ has a convergent subsequence. The following lemma asserts that compact metric spaces enjoy a similar property:

Lemma 2.5.7. *Let (K,d) be a compact metric space. Then every sequence in K has a convergent subsequence.*

Proof. Let $(x_n)_{n=1}^\infty$ be a sequence in K. Assume that $(x_n)_{n=1}^\infty$ has *no* convergent subsequence. This means that, for each $x \in X$ (it cannot be the limit of any subsequence of $(x_n)_{n=1}^\infty$!) there is $\epsilon_x > 0$ such that $B_{\epsilon_x}(x)$ contains only finitely many terms of $(x_n)_{n=1}^\infty$; that is, there is $n_x \in \mathbb{N}$ such that $x_n \notin B_{\epsilon_x}(x)$ for $n \geq n_x$. Since $\{B_{\epsilon_x}(x) : x \in K\}$ is an open cover for K, there are $x_1', \ldots, x_m' \in K$ with

$$K = B_{\epsilon_{x_1'}}(x_1') \cup \cdots \cup B_{\epsilon_{x_m'}}(x_m').$$

For $n \geq \max\{n_{x_1'}, \ldots, n_{x_m'}\}$, this means that

$$x_n \notin B_{\epsilon_{x_1'}}(x_1') \cup \cdots \cup B_{\epsilon_{x_m'}}(x_m') = K,$$

which is absurd. □

Proposition 2.5.8. *Let (K,d) be a compact metric space. Then K is both complete and separable.*

Proof. Let $(x_n)_{n=1}^\infty$ be a Cauchy sequence in K. By Lemma 2.5.7, $(x_n)_{n=1}^\infty$ has a convergent subsequence, say $(x_{n_k})_{k=1}^\infty$, whose limit we denote by x. Let $\epsilon > 0$. Then there is $k_\epsilon \in \mathbb{N}$ such that $d(x_{n_k}, x) < \frac{\epsilon}{2}$ for $k \geq k_\epsilon$. Furthermore, there is $n_\epsilon \in \mathbb{N}$ with $d(x_n, x_m) < \frac{\epsilon}{2}$ for $n \geq n_\epsilon$. Choose $k_0 \geq k_\epsilon$ so large that $n_{k_0} \geq n_\epsilon$. For $n \geq n_\epsilon$, we obtain that

$$d(x_n, x) \leq d(x_n, x_{n_{k_0}}) + d(x_{n_{k_0}}, x) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

It follows that $x = \lim_{n\to\infty} x_n$.

To see that K is separable, first note that $\left\{B_{\frac{1}{n}}(x) : x \in K\right\}$, the collection of all open balls in K of radius $\frac{1}{n}$, is an open cover for K for each $n \in \mathbb{N}$. Since K is compact, each such open cover has a finite subcover: there are, for each $n \in \mathbb{N}$, a positive integer m_n as well as $x_{1,n}, \ldots, x_{m_n,n} \in K$ such that

$$K = B_{\frac{1}{n}}(x_{1,n}) \cup \cdots \cup B_{\frac{1}{n}}(x_{m_n,n}).$$

The set $\bigcup_{n=1}^\infty \{x_{1,n}, \ldots, x_{m_n,n}\}$ is clearly countable. We claim that it is also dense in K. To see this, let $x \in K$, and let $\epsilon > 0$. Let $n \in \mathbb{N}$ be so large that $\frac{1}{n} < \epsilon$. Since $K = B_{\frac{1}{n}}(x_{1,n}) \cup \cdots \cup B_{\frac{1}{n}}(x_{m_n,n})$, there is $j \in \{1, \ldots, m_n\}$ such that $x \in B_{\frac{1}{n}}(x_{j,n})$ and thus $x_{j,n} \in B_\epsilon(x)$. □

We now turn to two notions related to compactness.

Definition 2.5.9. *Let (X,d) be a metric space. Then:*

(a) X *is called* totally bounded *if, for each* $\epsilon > 0$*, there are* $x_1, \ldots, x_n \in X$ *with*

$$X = B_\epsilon(x_1) \cup \cdots \cup B_\epsilon(x_n).$$

(b) X *is called* sequentially compact *if every sequence in* X *has a convergent subsequence.*

Some relations among compactness, total boundedness, and sequential compactness are straightforward. Every compact metric space is trivially totally bounded and also sequentially compact by Lemma 2.5.7. On the other hand, $(0,1)$ is easily seen to be totally bounded, but fails to be compact. The following theorem relates compactness, total boundedness, and sequential compactness in the best possible manner.

Theorem 2.5.10. *The following are equivalent for a metric space* (X,d)*.*

(i) X *is compact.*
(ii) X *is complete and totally bounded.*
(iii) X *is sequentially compact.*

Proof. By Lemma 2.5.7, (i) $\Longrightarrow$ (iii) holds.

(iii) $\Longrightarrow$ (ii): The same argument as in the proof of Proposition 2.5.8 shows that X is complete. Assume that X is not totally bounded. Then there is ϵ_0 such that

$$B_{\epsilon_0}(x_1') \cup \cdots \cup B_{\epsilon_0}(x_n') \subsetneq X$$

for any choice of $x_1', \ldots, x_n' \in X$. We use this to inductively construct a sequence in X that has no convergent subsequence. Let $x_1 \in X$ be arbitrary. Pick $x_2 \in X \setminus B_{\epsilon_0}(x_1)$. Then pick $x_3 \in X \setminus (B_{\epsilon_0}(x_1) \cup B_{\epsilon_0}(x_2))$. Continuing in this fashion, we obtain a sequence $(x_n)_{n=1}^\infty$ in X with

$$x_{n+1} \notin B_{\epsilon_0}(x_1) \cup \cdots \cup B_{\epsilon_0}(x_n) \qquad (n \in \mathbb{N}).$$

It is clear from this construction that

$$d(x_n, x_m) \geq \epsilon_0 \qquad (n, m \in \mathbb{N},\, n \neq m),$$

so that no subsequence of $(x_n)_{n=1}^\infty$ can be a Cauchy sequence. This is impossible if X is sequentially compact.

(ii) $\Longrightarrow$ (i): Let $\mathcal{U}$ be an open cover of X, and assume that it has no finite subcover. Since X is totally bounded, it can be covered by finitely many open balls of radius 1. Consequently, there is at least one $x_1 \in X$ such that $B_1(x_1)$ cannot be covered by finitely many sets from $\mathcal{U}$. Again by the total boundedness of X, the open ball $B_1(x_1)$ can be covered by finitely many open balls of radius $\frac{1}{2}$ (not necessarily centered at points of $B_1(x_1)$). Consequently, there is at least one $x_2 \in X$ such that $B_{\frac{1}{2}}(x_2) \cap B_1(x_1)$ cannot be covered by finitely many sets from $\mathcal{U}$. Continuing this construction, we obtain a sequence $(x_n)_{n=1}^\infty$ in X such that

$$B_{\frac{1}{n}}(x_n) \cap \cdots \cap B_{\frac{1}{2}}(x_2) \cap B_1(x_1)$$

cannot be covered by finitely many sets from $\mathcal{U}$. For $n \in \mathbb{N}$, let

$$F_n := \overline{B_{\frac{1}{n}}(x_n) \cap \cdots \cap B_{\frac{1}{2}}(x_2) \cap B_1(x_1)}.$$

Since $\operatorname{diam}(F_n) \leq \frac{2}{n} \to 0$, Cantor's intersection theorem yields that $\bigcap_{n=1}^{\infty} F_n = \{x\}$ for some $x \in X$. Let $U_0 \in \mathcal{U}$ be such that $x \in U_0$, and let $\epsilon > 0$ be such that $B_\epsilon(x) \subset U_0$. Choose $n_\epsilon \in \mathbb{N}$ such that $\frac{2}{n_\epsilon} < \epsilon$. Since $\operatorname{diam}(F_{n_\epsilon}) \leq \frac{2}{n_\epsilon}$, this means that $F_{n_\epsilon} \subset B_\epsilon(x) \subset U_0$. In particular, $\{U_0\}$ is a finite cover of $B_{\frac{1}{n_\epsilon}}(x_{n_\epsilon}) \cap \cdots \cap B_1(x_1)$, which is impossible according to our construction. □

Corollary 2.5.11. *Let (X, d) be a totally bounded metric space. Then its completion is compact.*

Proof. Let $\left(\tilde{X}, \tilde{d}\right)$ be the completion of (X, d). For $r > 0$ and $x \in X$, we write $B_r(x; X)$ and $B_r\left(x; \tilde{X}\right)$ for the open balls with radius r centered at x in X and $\tilde{X}$, respectively.

Let $\epsilon > 0$. Since X is totally bounded, there are $x_1, \ldots, x_n \in X$ such that

$$X = B_{\frac{\epsilon}{2}}(x_1; X) \cup \cdots \cup B_{\frac{\epsilon}{2}}(x_n; X) \subset B_{\frac{\epsilon}{2}}\left(x_1; \tilde{X}\right) \cup \cdots \cup B_{\frac{\epsilon}{2}}\left(x_n; \tilde{X}\right).$$

Now, $\overline{B_{\frac{\epsilon}{2}}\left(x_1; \tilde{X}\right)} \cup \cdots \cup \overline{B_{\frac{\epsilon}{2}}\left(x_n; \tilde{X}\right)}$ is a closed subset of $\tilde{X}$ containing X and therefore must be all of $\tilde{X}$. Since $\overline{B_{\frac{\epsilon}{2}}\left(x_j; \tilde{X}\right)} \subset B_\epsilon\left(x_j; \tilde{X}\right)$ for $j = 1, \ldots, n$, we obtain that

$$\tilde{X} = B_\epsilon\left(x_1; \tilde{X}\right) \cup \cdots \cup B_\epsilon\left(x_n; \tilde{X}\right).$$

Hence, $\tilde{X}$ is also totally bounded and thus compact by Theorem 2.5.10. □

The Heine–Borel theorem, which characterizes the compact subsets of $\mathbb{R}^n$, is probably familiar from several variable calculus. At the end of this section, we deduce it from Theorem 2.5.10, thus increasing our stock of compact and noncompact metric spaces.

Corollary 2.5.12 (Heine–Borel theorem). *Let $K \subset \mathbb{R}^n$. Then K is compact if and only if it is bounded and closed in $\mathbb{R}^n$.*

Proof. In view of Example 2.5.3(b) and Proposition 2.5.4(ii), the "only if" part is clear.

For the converse, first note that, since K is bounded, there is $r > 0$ such that $K \subset [-r, r]^n$. Since K is closed in $\mathbb{R}^n$ and therefore in $[-r, r]^n$, we can invoke Proposition 2.5.4(i) and suppose without loss of generality that $K = [-r, r]^n$.

As a closed subset of a complete metric space, K is clearly complete. It is therefore sufficient to show that K is totally bounded. Let $\epsilon > 0$. For $m \in \mathbb{N}$ and $j \in \{1, \ldots, m\}$, let

$$I_j := \left[-r + (j-1)\frac{2r}{m}, -r + j\frac{2r}{m}\right],$$

and note that

$$[-r, r] = \bigcup_{j=1}^{m} I_j$$

and thus

$$K = \bigcup_{(j_1,\ldots,j_n)\in\{1,\ldots,m\}^n} I_{j_1} \times \cdots \times I_{j_n}.$$

Let $(j_1, \ldots, j_n) \in \{1, \ldots, m\}^n$, and let $x, y \in I_{j_1} \times \cdots \times I_{j_n}$. The Euclidean distance of x and y can then be estimated via

$$\|x - y\| = \sqrt{\sum_{k=1}^{n}(x_k - y_k)^2} \leq \sqrt{\sum_{k=1}^{n}\left(\frac{2r}{m}\right)^2} = \frac{2r}{m}\sqrt{n}.$$

Let m be so large that $\frac{2r}{m}\sqrt{n} < \epsilon$. For $(j_1, \ldots, j_n) \in \{1, \ldots, m\}^n$, let $x_{(j_1,\ldots,j_n)} \in I_{j_1} \times \cdots \times I_{j_n}$. By the foregoing estimate, $I_{j_1} \times \cdots \times I_{j_n} \subset B_\epsilon\left(x_{(j_1,\ldots,j_n)}\right)$ holds, so that

$$K \subset \bigcup_{(j_1,\ldots,j_n)\in\{1,\ldots,m\}^n} B_\epsilon\left(x_{(j_1,\ldots,j_n)}\right).$$

Consequently, K is totally bounded and therefore compact. □

Outside the realm of Euclidean n-space, the Heine–Borel theorem is no longer true, and even worse: for general metric spaces, it fails to make sense. First of all, every metric space is closed in itself, so that requiring a set to be closed depends very much on the metric space in which we are considering it. Secondly, what does it mean for a subset of a metric space to be bounded? We could, of course, define a set to be bounded if it has finite diameter, but since *every* metric is equivalent to a metric that attains its values in $[0, 1)$, and since compactness is not characterized via a particular metric, but rather through open sets, boundedness cannot be used in general metric spaces to characterize compactness.

In normed spaces, it still makes sense to speak of bounded sets as in $\mathbb{R}^n$, but the Heine–Borel theorem becomes false.

Example 2.5.13. Let $E = C([0,1], \mathbb{F})$ be equipped with $\|\cdot\|_\infty$, and let $(f_n)_{n=1}^\infty$ be defined by

$$f_n\colon [0,1] \to \mathbb{R}, \quad t \mapsto t^n \qquad (n \in \mathbb{N}).$$

This sequence is contained in the closed unit ball $B_1[0]$ of E. If the Heine–Borel theorem is true for E, then $B_1[0]$ is compact and, consequently, $(f_n)_{n=1}^\infty$ has a convergent subsequence, say $(f_{n_k})_{k=1}^\infty$, with limit f. Since $f_n \in S_1[0]$ for $n \in \mathbb{N}$ and since $S_1[0]$ is closed, it is clear that $f \in S_1[0]$; that is, $\|f\|_\infty = 1$. On the other hand, convergence in E is uniform convergence and thus entails pointwise convergence. Hence, we have for $t \in [0,1)$ that

$$f(t) = \lim_{k\to\infty} f_{n_k}(t) = \lim_{k\to\infty} t^{n_k} = 0.$$

Since f is continuous, this means that $f(1) = 0$ as well and thus $f \equiv 0$. This is a contradiction.

More generally, the closed unit ball of a normed space E is compact if and only if $\dim E < \infty$ (Theorem B.5).

Exercises

1. Show that a discrete metric space (X, d) is compact if and only if X is finite.
2. Let (X, d) be a metric space, and let $(x_n)_{n=1}^\infty$ be a sequence in X with limit x_0. Show that the subset $\{x_0, x_1, x_2, \ldots\}$ of X is compact.
3. Let (X, d) be a metric space, and let F and K be subspaces of X such that F is closed in X and K is compact. Show that
$$F \cap K \neq \varnothing \quad \Longleftrightarrow \quad \inf\{d(x,y) : x \in F,\ y \in K\} = 0.$$
What happens if we replace the compactness of K by the demand that it be closed in X?
4. Let $(K_1, d_1), \ldots, (K_n, d_n)$ be compact metric spaces, and let $K := K_1 \times \cdots \times K_n$ be equipped with any of the two (equivalent) metrics D_1 and D_∞ from Example 2.3.13(b). Show that K is compact.
5. More generally, let $((K_n, d_n))_{n=1}^\infty$ be a sequence of compact metric spaces, and let $K := \prod_{n=1}^\infty K_n$ be equipped with a metric d as in Example 2.1.2(h). Show that (K, d) is compact.
6. Let E be a normed space, and let $K, L \subset E$ be compact. Show that $K + L := \{x + y : x \in K,\ y \in L\}$ is also compact.
7. A subset S of a metric space (X, d) is called *relatively compact* if $\overline{S}$ is compact. Show that $S \subset X$ is relatively compact if and only if each sequence in S has a subsequence that converges in X. To what more familiar notion is relative compactness equivalent if the surrounding space X is complete?
8. Let (X, d_X) and (Y, d_Y) be metric spaces. A function $f : X \to Y$ is called *uniformly continuous* if, for each $\epsilon > 0$, there is $\delta > 0$ such that $d_Y(f(x), f(y)) < \epsilon$ whenever $d_X(x, y) < \delta$. Show that any continuous function from a compact metric space into another metric space is uniformly continuous.
9. *Lebesgue's covering lemma*. Let (K, d) be a compact metric space, and let $\mathcal{U}$ be an open cover of K. Show that there is a number $L(\mathcal{U}) > 0$ (the *Lebesgue number* of $\mathcal{U}$) such that any $\varnothing \neq S \subset K$ with $\operatorname{diam}(S) < L(\mathcal{U})$ is contained in some $U \in \mathcal{U}$.

Remarks

Metric spaces are little more than one hundred years old: their axioms appear for the first time in Maurice Fréchet's thesis [Fréchet 06] from 1906. Instead of metric spaces, Fréchet speaks of classes (E), and the distance of two elements with respect to the given metric is called their *écart*, which is French for *gap*. A few years later, the German mathematician Felix Hausdorff rechristened Fréchet's classes (E) in his treatise [Hausdorff 14]: he called them *metrische Räume*, which translates into English literally as *metric spaces*. Most of the material from Sections 2.1, 2.2, 2.3, and 2.5 can already be found in [Hausdorff 14].

What we call a semimetric is usually called a *pseudometric*. However, a map p from a linear space into $[0, \infty)$ that satisfies all the axioms of a norm, except that it allows that $p(x) = 0$ for nonzero x, is called a *semi*norm, not a *pseudo*norm. This is our reason for deviating from the standard terminology, so that $p(x - y)$ for a *semi*norm p defines a *semi*metric, which is a metric if and only if p is a norm.

Bourbaki's Mittag-Leffler theorem (Theorem 2.4.14) is from "his" monumental treatise *Eléments de mathématique* [Bourbaki 60]. The possessive pronoun is in quotation marks because Nicolas Bourbaki is not one man but the collective pseudonym of a group of French mathematicians that formed in 1935 and, from 1939 on, started publishing the aforementioned multivolume opus *Eléments de mathématique* with the goal to rebuild mathematics from scratch. Members of Nicolas Bourbaki have to leave once they reach age 50, and new members are appointed to replace the retiring ones. Hence, Nicolas Bourbaki is a truly immortal mathematician! Even though it is widely claimed (and believed), Nicolas Bourbaki was *not* the name of a French general in the Franco–Prussian war of 1871: there was a general in that war by the last name of Bourbaki, but his first names were Charles Denis. (He was offered the throne of Greece in 1862, which he turned down, and in the Franco–Prussian war, he unsuccessfully attempted suicide in order to avoid the humiliation of surrender.)

For a good reason, our Theorem 2.4.14 is somewhat less general than the result from [Bourbaki 60]. As Jean Esterle remarks in [Esterle 84]:

> Incidentally, the reader interested in a French way of writing a result as clear as Corollary 2.2 [$\approx$ Theorem 2.4.14] in a form almost inaccessible to human mind is referred to the statement by Bourbaki [...].

In statement and proof of Theorem 2.4.14, we follow [Dales 78].

Baire's theorem is sometimes referred to as *Baire's category theorem* (especially in older books). The reasons for this are historical. A subset of a metric space is called *nowhere dense* if its closure has an empty interior. A subset that is a countable union of nowhere dense sets used to be called a set of the *first category* in the space, and all other subsets were said to be of the *second category*. In this terminology, Baire's theorem (or rather Corollary

2.4.17) asserts that every complete metric space is of the second category in itself. The first/second category terminology has not withstood the test of time (when mathematicians nowadays speak of categories, they mean something completely different), but the nametag *category theorem* still survives to this day.

Maurice Fréchet died in 1973, at the age of 94, decades after the concept he had introduced in his thesis had become a mathematical household item.

http://www.springer.com/978-0-387-25790-7

A Taste of Topology
Runde, V.
2005, X, 182 p. 17 illus., Softcover
ISBN: 978-0-387-25790-7